

CS 3001 Notes

CS Ethics Notes Based On Lectures/Lecture Slides

Krish Katariya

Last updated: **February 27, 2024**

Contents

1. ADD HOW TO ARGUE SECTION	1
2. Therac-25	1
2.1. Software Errors	1
2.2. Design Flaws	2
2.3. Moral Responsibility	2
3. Utilitarianism	2
4. Kantianism / Deontology	3
4.1. Duties	3
4.2. Imperatives	3
5. Social Contract	3
5.1. Hobbes' Beliefs	3
5.2. Social Contract Theory	4
5.3. Rights	4
6. Stakeholder Analysis	4
7. Kohlberg's Stages of Moral Development	5
7.1. Stages 1 & 2	5
7.2. Stages 3 & 4	5
7.3. Stage 5 & 6	5
8. Virtue Ethics	5
8.1. Censorship	6
8.2. Freedom of Expression	6
9. Professional Ethics	7
9.1. ACM Code of Ethics	7
10. Internal Realism	7
11. Evidentialism	7
12. Reliabilism	8

1. ADD HOW TO ARGUE SECTION

2. Therac-25

Definition 2.1

The **therac-25** was a linear accelerator created to treat cancer with radiation therapy

The therac-25 was an all new version of earlier therac-6 and therac-20 machines, but unlike them did not have manual overrides.

The therac-25 had malfunctions and would give too much, or too little (which could also be dangerous) radiation causing lots of harm.

2.1. Software Errors

Two of the therac-25's main software errors were **race conditions**.

Definition 2.1.1

Race Conditions is when a system tries to perform two operations at very similar times but because of the nature of the system, the operations have to be done in specific manner for it to work correctly.

One race condition was with the system acting on old information that had been changed to new information while the computer wasn't checking the values.

2.2. Design Flaws

Design process flaws by reusing older code designed to work with a system that had manual overrides and locks.

- So the flaws were system flaws, not only software errors

The system was not designed with being fail safe in mind

2.3. Moral Responsibility

In order for a moral agent to be responsible for a harmful event, two conditions must hold:

- **Casual Condition:** The actions (or inactions) of agents must have caused the harm
- **Mental Condition:** The actions (or inactions) of agents must have been intended or willed by the agent*

*This also includes unintended harm if it came from carelessness or negligence

3. Utilitarianism

Theorem 3.1

In utilitarianism, an action is good if the benefits exceed its harms, and an action is bad if its harms exceed its benefits

Definition 3.1

Act Utilitarianism is the ethical theory that an action is good if the net effect (over anyone and everyone who is affected) is to produce more happiness than unhappiness

Case For:

- It focuses on happiness
- It is practical
- It is generally comprehensive

Case Against:

- Where to draw line in calculations? How bad is too bad if the end justifies the means?
- Takes a lot of time for each moral calculation
- Ignores innate sense of

Definition 3.2

Rule Utilitarianism is the ethical theory that holds that we ought to adopt those moral rules that, if followed by everyone, lead to the greatest increase in total happiness over all affected parties

A *rule utilitarian* chooses to follow a moral rule because its universal adoption would result in the greatest net increase in happiness. *This differs from kantianism because a rule utilitarian is still concerned with the consequences and not the motive.*

Case For:

- Not every moral decision requires performing utilitarian calculus
- Exceptional situations don't overthrow moral rules

- Reduces problem of **bias**(?)
 - idk what this means

Case Against:

- It forces us to use a single scale or measure to evaluate completely different kinds of consequences
- It ignores the problem of an unjust distribution of good consequences

4. Kantianism / Deontology

Unlike utilitarianism, Kantianism is not concerned with outcomes. Instead, the motive for why you did certain actions is more important.

4.1. Duties

Definition 4.1.1

Perfect duties are duties we are obligated to fulfill in every instance. **Imperfect duties** we are obligated to fulfill in general but not in every instance.

4.2. Imperatives

An **imperative** is a way in which reason commands will.

Definition 4.2.1

A **hypothetical imperative** is a conditional rule of the form "If you want X then do Y" A **categorical imperative** is an unconditional rule: a rule that always applies, regardless of circumstance. For Kant, only a categorical imperative can be a moral imperative.

Both of the *main two categorical imperatives* should be held for an act to be considered morally right.

1. Act only from moral rules that you can at the same time will to be universal laws
2. Act so that you always treat both yourself and other people as ends in themselves, and never only as a means to an end

The Case For:

- It treats all person as moral equals
- It gives all persons moral worth by considering them rational, autonomous beings
- Everyone is held to the same standard
- Kantianism produces universal moral guidelines

The Case Against:

- Sometimes no single rule can fully characterize an action
- Sometimes there is no way to resolve a conflict between rules
- Kantianism allows no exceptions to perfect duties

5. Social Contract

5.1. Hobbes' Beliefs

Man is inherently *evil*. In order to avoid man's evil state of nature, rational people understand that some sort of social cooperation is only possible when people mutually agree to follow certain guidelines.

Theorem 5.1.1

This leads to a **social contract** - the view that persons' moral and/or political obligations are dependent upon a contract or agreement among them to form the society in which they live.

5.2. Social Contract Theory

Theorem 5.2.1

Hobbes argues that everyone in society has implicitly agreed to two things:

1. The establishment of a set of moral rules to govern relations among citizens
2. A government capable of enforcing these rules

Hobbes believes that the morals derived from natural law do not permit individuals to challenge the laws of the sovereign, meaning that you cannot disobey the laws of the government.

5.3. Rights

Rights can be classified according to the duties they put on others

Definition 5.3.1

Negative Rights are rights that another can guarantee by leaving you alone to exercise your right.

- Right of free speech
- Right to be safe
- Right to move around

Positive Right - a right that obligates others to do something on your behalf (guarantee of an action)

- Right to free education
- Community and national security

Absolute Right - guaranteed without exception **Limited Right** - may/can be restricted

The Case For:

- It is framed in the language of rights
- It is based on a solid understanding of human nature, recognizing that rational people act out of self-interest in the absence of a common agreement
- It explains why under certain circumstances civil disobedience can be the morally right decision

The Case Against:

- None of us signed the social contract
- Some actions can be characterized in multiple ways

How do the philosophies compare? Reason for making a rule:

- **Utilitarianism**: it causes the greatest happiness for the most people
- **Kantianism**: it could be universalized for all people
- **Social contract theory**: rational people would collectively accept it as binding because of its benefits to the community

6. Stakeholder Analysis

Definition 6.1

Stakeholders are interested parties: people or entities that will be affected in a given situation

Make a list of all the stakeholders involved, and try to balance the positive and negative impact on people.

This is not an ethical theory, but it is a *useful way of looking at things*.

7. Kohlberg's Stages of Moral Development

7.1. Stages 1 & 2

Egocentric understanding of fairness based on individual need (kids)

1. Obedience and punishment orientation
2. Self interest orientation

7.2. Stages 3 & 4

Shared concept of fairness based in societal agreement

1. Interpersonal accord and conformity (good boy/girl attitude)
2. Authority and social-order maintaining (law and order morality)

7.3. Stage 5 & 6

Free standing logic of equality and reciprocity

1. Social Contract Orientation
2. Universal Ethical Principles (Kant)

8. Virtue Ethics

Virtue ethics can address the questions of how one should live, what kind of person one should become, and even what one should do in certain scenarios. Would a good person perform this action under this virtue?

Definition 8.1

Virtues: behaviors showing high standards. The qualities or states of character that find expression in morally good actions and morally good purposes. **Intellectual Virtues** - virtues associated with reasoning and truth **Moral Virtues** - habits or dispositions formed through the repetition of relevant virtuous actions. *A deep seated character trait.* Not simple just a disposition to act a certain way but also to *feel* that way.

Definition 8.2

Vices: the opposite from virtues. A vice is a character trait that prevents a human being from flourishing or being truly happy

Usually, there are two vices for every virtue, because one corresponds to excess and the other corresponds to a deficiency.

Ex: Rashness ← Courage → Cowardice

Note 8.1

Virtue ethics pays particular attention to the agent as well as the action and the consequences of the action. This touches on social contract theory, utilitarianism, and kantianism. *A good person does the right thing at the right time for the right reason.*

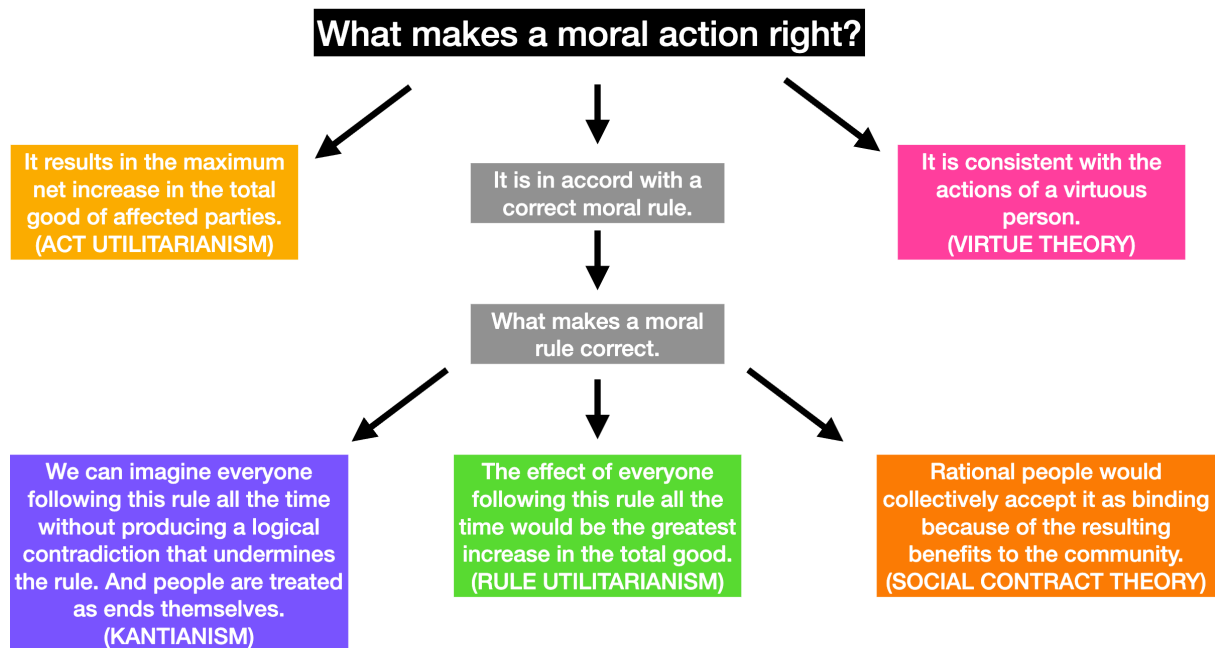
Pros

- In many situations it makes more sense to focus on virtues than on obligations or rights
- Personal relationships can be morally relevant to decision making
- There are no irresolvable moral dilemmas
- Recognizes the important of having a say in living a moral life

Cons

- Different people may have quite different conceptions of what human flourishing is
- Virtue ethics cannot be used to guide government policy

- Undermines attempts to hold people responsible for their bad actions



8.1. Censorship

Direct Censorship

- **Government monopolization** - government owns TV stations, radio, and newspapers. This has become harder due to the internet.
- **Prepublication review** - most governments have laws restricting the publication of information that harm national security
- **Licensing and Registration** - begins as a means to control media with limited network bandwidth (radio station)
 - However, controlling bodies such as FCC also censor or ban content on platforms wishing to obtain licenses

Self Censorship

- A group deciding for itself not to publish content for whatever reason.

Challenges posted by the internet

1. Unlike traditional 1-many media, the internet supports many to many communications
2. The internet is dynamic
3. The internet is huge
4. The internet is global
5. It is hard to distinguish between children and adults on the internet

8.2. Freedom of Expression

Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the government for redress of grievances

Note 8.2.1

Freedom of expression is not an absolute right.

Protection is not given to “libel, reckless or calculated lies, slander, misrepresentation, perjury, false advertising, obscenity and profanity, solicitation of crime, and personal abuse or fighting words”

9. Professional Ethics

Definition 9.1

Professional ethics refers to the moral principles that govern the behavior of individuals in a professional context.

How are Professional Ethics Different?

- Customers don't know the domain, so they rely on the honesty, knowledge, and expertise of the professional
- Professionals' products affect lots of people
 - People affected may be connected indirectly
 - Didn't have a chance to choose the professional
- Special responsibility not just to the customer, but to the public

9.1. ACM Code of Ethics

1. **Public** - Software Engineers shall act consistently with the public interest
2. **Client and Employer** - Software Engineers shall act in a manner that is in the best interests of their client and employer
3. **Product** - Software Engineers shall ensure that their products meet the highest possible standards possible
4. **Judgement** - Software Engineers shall maintain integrity and independence in a professional Judgement
5. **Management** - leaders shall subscribe to and promote an ethical approach to software development and maintainence
6. **Profession** - Software Engineers shall advance the integrity and reputation of the profession consistent with public interest
7. **Colleagues** - Software Engineers shall be fair to and supportive of their Colleagues
8. **Self** - Software Engineers shall participate in lifelong learning regarding the practice of their profession and shall promote an ethical approach to the profession

Definition 9.1.1

Exclusive Responsibility: “I am responsible for paying this bill” **Not exclusive:** “My boss is responsible for this, and so am I” Moral responsibility is typically not exclusive

10. Internal Realism

- How do we know there is a desk here?
 - We are all trapped in our subjective perceptions of the world around us, but our subjective perceptions oddly agree

11. Evidentialism

Definition 11.1

Evidentialism: all beliefs must be supported by evidence

- Some beliefs are basic and justified, such as things you learn through a reliable sense

12. Reliabilism

Definition 12.1

Reliabilism: a belief is reliable if it is the result of a reliable cognitive process.

- A *reliable* process is one that generally leads to the correct outcomes