

• Software to grant bail to ppl in jail / prison of America

Date \_\_\_\_\_  
Page \_\_\_\_\_

'Philosophy encompasses everything'  
-literal statement

final degree you get in anything is Phd

Science is also  
Philosophy!

↓  
Doctor in  
Philosophy

Basic Pillars of Philosophy:

Q 1) What is all this? → Ontology

↳ What is reality  
composed of?

Q 2) How do I know this right?

What is truth? What is knowledge?

→ Epistemology

Q 3) What do I need to do? → After first 2 are answered  
→ Ethics (Moral Philosophy)

\* ETHICS:

1) Normative ethics (What is the ethical thing to do?)

2) Meta ethics (What is ethics?)

3) Applied ethics (How do I transform these ethical ideas to reality?)

Oldest model of ethics that we <sup>know</sup> ~~are~~:

Virtue ethics → Does a person have good qualities (good)

Is the person virtuous

(Don't care about their actions / thinking)



- Deontological ethics: (our constitution)  
makes it all about the action, whether action is good or bad.

Advocated by Immanuel Kant

- Categorical imperative
- Veil of ignorance
- Natural rights

Regardless of who or why did it, some actions are bad.

John Rawls: wheel of ignorance

- To decide whether anything is wrong or right / good or bad, one should come like a total empty slate.

- Consequentialism: (our course)

Only look at what <sup>how</sup> the action impacts society. What is good is what makes ppl happy. (Utilitarianism) = Given by Bentham.

- State consequentialism

Karl Marx called this shopkeeper mentality.  
↳ Opposed state consequentialism

- Contextual

- Welfareism

- Pragmatism:

(Believes on representativeness)

American philosophy of ethics.

Get everybody of diff castes & classes on a table & find a sol<sup>n</sup> to which everyone agrees.

India constitution is based on this.

B-R Ambedkar studied under American prof.



## \* Applied ethics

- MIT machine ethics experiment  
→ suppose a car on road has to kill x or y, which one should be killed?
- Policymaking
- Bioethics & Machine ethics
- AI control problem  
(we should be governing AI not the other way around)  
Chess example
- Effective altruism, Longtermism  
Example: Sam Bankman
- Is political-economy truly amoral?  
→ Economic justice cannot come at the cost of social justice



meta ethics:

What is the meaning of moral statements?

Are moral statements an attempt to describe reality? (Truth apt statements)

Is it similar to descriptive statement?

People who believe this are called cognitivists.  
(Moral statements <sup>have</sup> true / false answers)

Other are called non-cognitivists.

View of cognitivist:

Are only moral statements true?

→ No, all are false, values don't exist: Error Theory

→ Yes, some moral statements are true:

Subjective: Ethical subjectivism / relativism

Objective: Moral realism - mind independent morality exists.

Moral  
neutrality

Moral  
non-neutrality

View of non cognitivist:

→ What are we doing if not fact like statements?

• We are making emotions: Emotivism

• Quasi-realism: Morality not an actual, but a pseudo description which we have made to live our life

• We are making commands: Prescriptivism

(Cognitive → Descriptivists → non-prescriptivists)

we pretend  
they are  
real



→ Is ought gap — David Hurst  
Most moral arguments start with logical statements, but suddenly at 'ought' comes  
— Description change to Prescription.

Where does the reality 'should' be like something  
come from?

Keep description & desire separate when you  
think of policy.

→ Ethics :

IEEE, ACM, UN started taking ethics  
initiatives for AI (Recently around 2016)

Companies started establishing ethics boards  
for future proofing regarding AI in their  
own companies.

• Project Maven

we will often make a mistake of considering  
a company as a monolith.

Idea of Pentagon

we want the bio data of any person any  
where at any time — Full history of every  
person.

Pentagon gave this project to Google.

Engineers at Google said we won't do this.

they went on a strike after understanding  
the implication of the project.



Page 15

Date \_\_\_\_\_  
Page \_\_\_\_\_

Regulation comes from the top but its the bottom which make the AI

- How ethics get decided is very non-transparent  
Indian standards organization is not democratic not open for all to join  
PPI who end up making these standards -  
their selection & also std making thinking etc  
is not very transparent



## 2<sup>nd</sup> Topic From Philosophy

### WHAT IS A MIND?

What causes consciousness? Is the mind

Nobody knows what consciousness is

Chinese room - from room only AI sheet can be sent out or sent in

Room has a magic dictionary & a person who doesn't know anything about Chinese

→ Similarly AI itself doesn't know it is intelligent

→ Neural network is just a translator who follows given code to give an output

Mind internally knows 'I exist' that's why we are not a computer

\* Dualism: Mind & body are separate (Aatma)  
Soul

Property dualism

Substance dualism

Monism / Idealism: mind can only perceive ideas

Materialism: Using mind it is possible to measure reality outside of mind

Physics <sup>governs</sup> Matter <sup>makes</sup> Mind

Then everything was pre-planned by Physics from Big Bang?



En  
Paul S. (1-11)

Date \_\_\_\_\_  
Page \_\_\_\_\_

Lobotomy - Drill hole in brain to change behaviour of a person.

Go through the slides for remaining concepts