

U3APP 2019 Term1

Philosophy and the challenges of the 21st century

Class 2_12 February 2019: Artificial Intelligence (AI) (continued)

The questions in 2, 3 and 4 below are designed to direct your thinking as you listen to/read/view the resource material. Further optional resources are listed in 5. And you should feel free to explore the internet for additional material.

If time permits, we will introduce the topic of Predictive Policing. However main discussion will be held over until Class 3.

1. General questions from last class concerning the nature of “intelligence”, “learning”, etc.

We will look briefly at any questions raised by last week’s discussion

2. Autonomous Agents and super AI

Listen to [What intelligent machines tell us about ourselves](https://www.abc.net.au/radionational/programs/godforbid/), “God Forbid” ABC Radio National, Sunday 3 Feb 2019 6am: <https://www.abc.net.au/radionational/programs/godforbid/>

Should we be concerned about claims that AI programs are capable of learning? Consider what is meant by “learning”: Can we separate learning cognitive skills (e.g., maths, reading, etc) from learning how to be ethical? What makes the difference? Is consciousness/awareness necessary for learning?

Some of our learning is implicit, and comes out in the form of bias. Can AI be biased? Machines can develop abilities to make discriminations much more finely than we humans can. Does this create ethical problems for such machines?

Do you think it is possible that intelligent machines might learn to outsmart humans? Is it a development to be encouraged? To what extent?

(From *The Conversation* article attached to letter for Class 1): *Risks will emerge from the fact that super-intelligent systems will identify more efficient ways of doing things, concoct their own strategies for achieving goals, and even [develop goals of their own](#). Imagine these examples:*

- *an AI system tasked with preventing HIV decides to eradicate the problem by killing everybody who carries the disease, or one tasked with curing cancer decides to kill everybody who has any genetic predisposition for it*
- *an autonomous AI military drone decides the only way to guarantee an enemy target is destroyed is to wipe out an entire community*
- *an environmentally protective AGI decides the only way to slow or reverse climate change is to remove technologies and humans that induce it.*

For a sci-fi scenario with super AI see Ted talk by Sam Harris:

https://www.ted.com/talks/sam_harris_can_we_build_ai_without_losing_control_over_it

How useful are **Asimov’s Laws of Robotics**¹

3. Feelings, empathy, emotions

Do feelings and emotions play a part in our moral life and our formation as moral subjects? Or is morality just a matter of acting rationally and dispassionately?

Consider: empathy, trust, regret, disgust, shame. Other?

Apply this to some actual cases – e.g., remote controlled military drones:

Is drone warfare different from conventional warfare in an ethically relevant way? Can the very idea of a war fought by remotely controlled drones be judged to be unethical? Where does responsibility lie?

Robo wars (P.W.Singer): <http://www.abc.net.au/radionational/programs/philosopherszone/the-morality-of-robo-wars---p.w.-singer/3701744>

Do you think it possible for a robot to experience feelings (e.g., trust, empathy, shame, etc.)? Does this matter when it comes to inter-personal relationships?

Robots and feeling pain (video): <https://aeon.co/videos/pain-leads-to-empathy-and-self-preservation-should-we-make-robots-feel-it>
<http://www.ethics.org.au/on-ethics/blog/december-2017/bladerunner-westworld-sexbot-suffering>

4. Relations with Others

What are the ethical issues raised by robotic carers?

Do these same issues arise in the case of sex robots?

What, if any, qualities should a robot exhibit if we are to consider granting it human rights?

Citizenship? Does AI lead to a debasement of human values?

Aged care robots: <https://www.youtube.com/watch?v=yYFLVFni1to>

Robots and sex: https://abcmedia.akamaized.net/rn/podcast/2015/09/rnd_20150930_1842.mp3

Citizen robots: <https://www.youtube.com/watch?v=E8Ox6H64yu8>

5. Further Optional reading/viewing/listening:

AI and bias: <https://www.theage.com.au/business/consumer-affairs/rental-apps-could-increase-discrimination-against-tenants-choice-20181121-p50hf4.html?btis>

<https://www.wired.com/story/tim-oreilly-algorithms-have-already-gone-rogue/>

Robots and feeling pain: <https://aeon.co/videos/pain-leads-to-empathy-and-self-preservation-should-we-make-robots-feel-it>

Japan carer robots: <https://www.google.com/search?client=firefox-b-d&q=japanese+robot+carer>

Aged care robots (audio): <http://www.abc.net.au/radionational/programs/lifematters/2016-03-22/7256014>

AI and Sex: movies “Ex Machina” “Bladerunner” “Her” raise relevant issues.

Citizen robots <http://www.abc.net.au/radionational/programs/theminefield/can-there-be-a-place-for-robots-within-the-human-community/9104338>

ⁱ Asimov_ Laws of Robotics:

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 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**.