HIM Lecture 4

Anthropomorphism: Can and Should Robots Be People?

Joanna J. Bryson
University of Bath, United Kingdom
@j2bryson

Outline

- Definitions
- Strong anthropomorphism
- Expedient anthropomorphism (next lecture)

Definitions

for communicating this course

- Intelligence is doing the right thing at the right time (in a dynamic environment).
- Agents are any vector of change,
  - e.g. chemical agents.
- Moral agents are considered responsible for their actions by a society.
- Moral patients are considered the responsibility of a society's agents.
- Artificial Intelligence is intelligence deliberately built.

} Arguably, ethics is determined by and determines a society–a constantly renegotiated set of equilibria. Law is a part of ethics by this definition.

Basic legal question: Is there anything about intelligent technology that changes responsibility for the intentional act of creation?

Questions of Anthropomorphism

1. Strong anthropomorphism:
    1. Should we build AI in such a way that artefacts should be moral subjects?
    2. Is it inevitable that as artefacts become more intelligent that they are owed moral subjectivity?
2. Expedient anthropomorphism:
    1. Is AI that appears humanlike easier to use / more effective?
    2. Is it moral to make AI appear more humanlike than it is?

Questions of Machine Anthropomorphism

1. Strong anthropomorphism:
    1. Should we build AI in such a way that artefacts should be moral subjects?
    2. Is it inevitable that as artefacts become more intelligent that they are owed moral subjectivity?
2. Expedient anthropomorphism:
    1. Is AI that appears humanlike easier to use / more effective?
    2. Is it moral to make AI appear more humanlike than it is?

Anthropomorphism & Dehumanisation

- Humans routinely promote and reduce assessed social distance, in-group / out-group status.
- Psychologists have found they can manipulate this with a variety of visual auditory cues; stress and threat indicators.
- In its most extreme, human enemies are portrayed as subhuman or nonhuman in advance of applying lethal force e.g. war, pogrom, death penalty.
- In contrast, pets and artefacts can be promoted to humanlike status e.g. religious icon, inherited wealth.

Even statues in public spaces alter human behaviour.

This fact violates the intuitions of many who support AI personhood on the basis of intuitions. (But see Kant, HIM5 / Wednesday's lecture.)

Statue by Seward Johnson

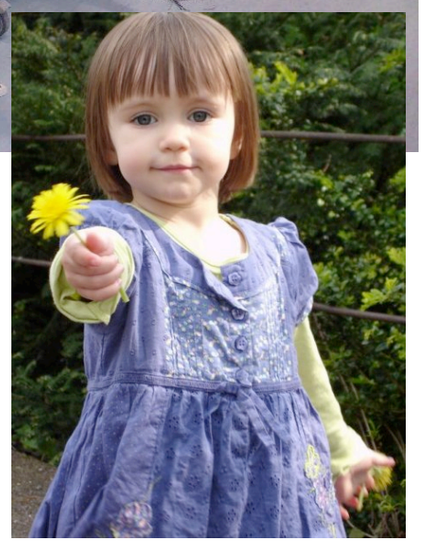
Forms of Personhood

- Biological persons (*Homo sapiens*)
  - May not be recognised as such e.g. race, gender, dead, absent & believed dead, comatose.
- Legal personhood
  - Includes corporations, icons (India), rivers (New Zealand).
- Moral personhood
  - Entities which arguably deserve the designation, e.g. chimpanzees, robots.

Outline

- Definitions
- Strong anthropomorphism
- Expedient anthropomorphism (next lecture)

There's no question whether we have the technical capacity to build synthetic legal persons.



(Bryson 2010, 2016, 2018)

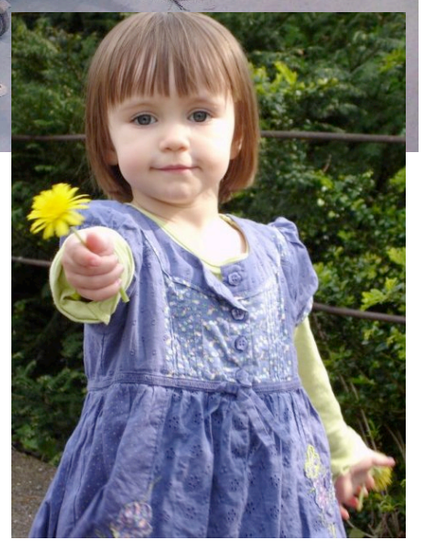
AI and law are both authored–cultural artefacts.
Science cannot determine AI's place in society–that decision is normative, not factual.
Science can predict outcomes of policy (inform such decisions).

photos: Georgio Metta (top) & Emmanuel Tanguy

The Study of Ethics: Moral Philosophy

- How do you determine an appropriate course of action? Normative Ethics
- What do people actually do? Descriptive Ethics
- How can we achieve moral outcomes? Applied Ethics
- Can ethics even make sense? Meta Ethics

Moral philosophy 101: Descriptive ≠ Normative. Is does not imply ought.
(But ought may imply or require can. But philosophers even argue about that!)



(Bryson 2010, 2016, 2018)

AI and law are both authored–cultural artefacts.
Science cannot determine AI's place in society–that decision is normative, not factual.
Science can predict outcomes of policy (inform such decisions).

photos: Georgio Metta (top) & Emmanuel Tanguy

Legal Personhood

1. Actual persons / citizens / landowners
   - (definition has been expanding)
   - organising coalitions via contracts.
2. Collections of humans, in order to simplify contracts and negotiation.
   - A fiction (hack) that only works because (or to the extent) corporations can be subjected to the same penalties as humans.

2. Collections of humans, in order to simplify contracts and negotiation.

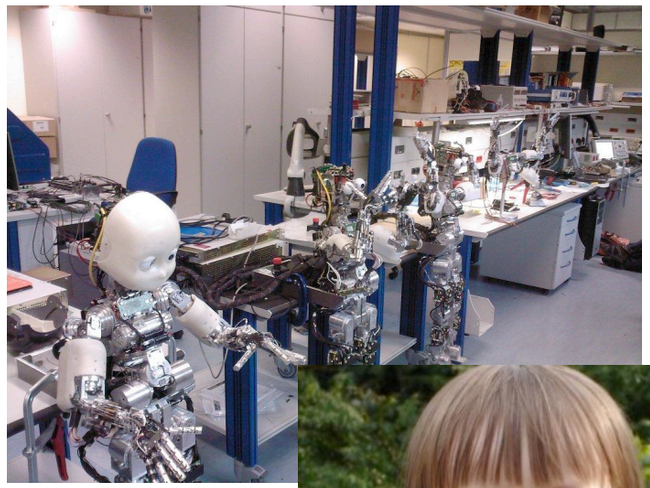
   - A fiction (hack) that only works because (or to the extent) corporations can be subjected to the same penalties as humans.

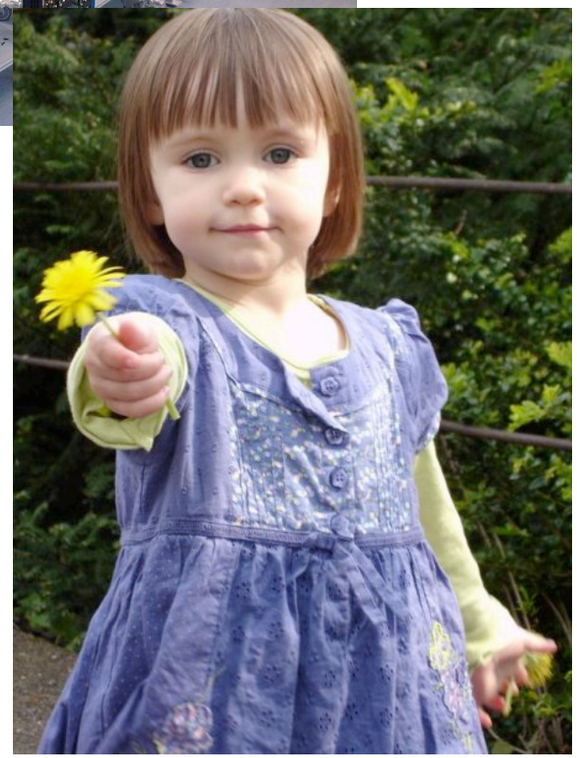
Fictitious Personhood

2. Collections of humans, in order to simplify contracts and negotiation.
    - A fiction (hack) that only works because (or to the extent) corporations can be subjected to the same penalties as humans.
- Overextended already (arguably).
- All the European Parliament really asked the European Commission to consider legislating for AI in 2017 (mostly about liability evasion for car manufacturers wrt driverless.)

Recompense

- Penalties in law have two purposes:
  - actual compensation
  - dissuasion.
- Folk psychology confounds these, but really jail time, fall in status, &c don't compensate.
- Implausible that built AI – designed & maintainable – will be subject to dissuasion.





(Bryson 2010, 2018)

Evolution assures that for any social species, social exclusion (or subjugation) is a severe disadvantage.

Authored systems are decomposable; legitimate (safe) products are modular and debuggable. Suffering in such is incoherent.

Liability is limited to money, plant, and data; no dissuasion is possible.

photos: Georgio Metta (top) & Emmanuel Tanguy

Real Persons

- Actual persons / citizens / landowners
  - (definition has been expanding)
  - organising coalitions via contracts.

- If we could build AI by cloning / whole brain uploading etc. that should probably be included in the expansion set.
  - But AI heaven isn't likely to be tractable, and no one thinks human cloning / ownership is ethical (Bryson 2010).

There's no question whether we have the technical capacity to build synthetic legal persons.

In my opinion, the real questions are:

Can we build a system we are not obliged to?

Are we obliged to do so if we can?

Can we build a system we are not obliged to?

Can we build a system we are not obliged to?

- Yes
- We already have (many times).
- We can eliminate non-replaceability by using mass-produced hardware and continuously backed-up memory.
- We can avoid resentment of subordinate position by not cloning evolved minds.

…at least in licensed commercial products.

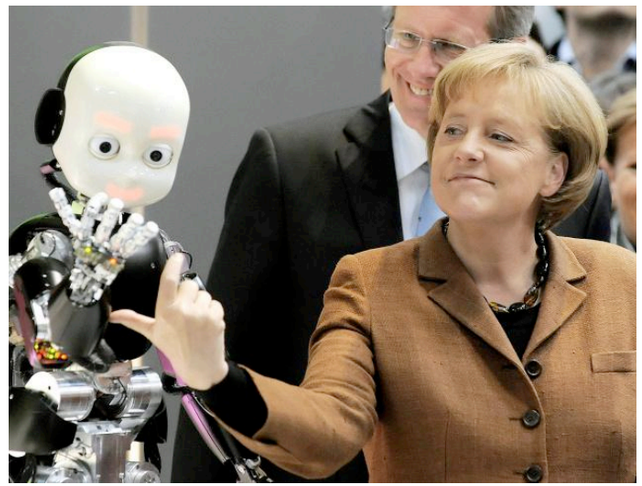
Are we obliged to do so if we can?

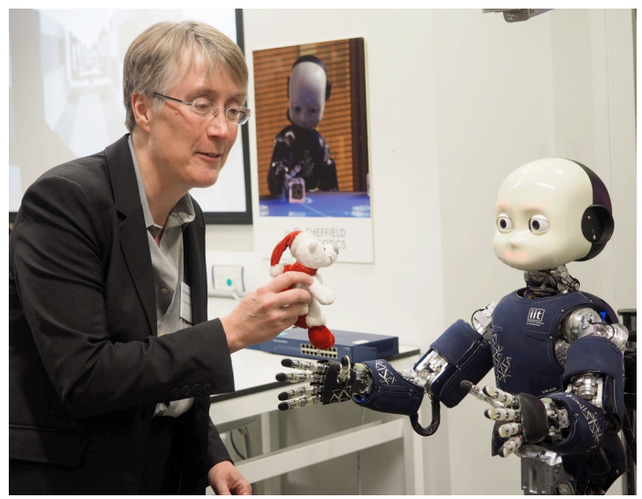
- Yes

Five Reasons Not to Other AI

#1 Moral Hazard

- We are preprogrammed to think humanoid robots are people (Kamewari &*al* 2005).
- So people will think we've made persons well before we have.
- Facilitates political and economic exploitation.





Bryson & Kime 1998, IJCAI 2011

#2 Second Order Moral Patiency

- Why should we build robots to suffer when they lose social status? To 'die' in fires? To mind being owned?
- We are obliged to build robots we are not obliged to.
- Not a double standard: pick one standard for moral subjects, don't build to it.

LF Miller 2015 Hellström 2013; Bryson 2016, 2007

#3 Fear of Robot Apocalypse Distracts from Real Threats

- AI is here now changing the world.
- By increasing communication, interdependence, discoverability, we decrease privacy and individual autonomy.
- Projecting AI into the future because "it's not human enough yet" endangers us now.

(Bryson 2015)

HM

Lecture 7

#4 Ethical Coherence

- What makes people special is that we're members of a social species – we've evolved in a context of interdependence(Zahavi 1977,Sylwester &*al* 2013).
- Society defines, enforces 'responsibility'; enforcement often through punishment (Solaiman 2016).
- Evolution ensures suffering, shame are inextricable parts of being human (also of apes, dogs).
- Good AI is modular; suffering in such is incoherent.
  - Clones should not be slaves, nor made.

#5 Legal Lacuna

HM

- Assigning responsibility / personhood to artefacts allows powerful individuals & organisations to avoid tax, legal liability.
  - Try suing a bankrupt robot.
- Already a problem: shell organisations (AI, cf. List & Pettit 2011) shield rich companies.
- One nightmare: Autocrats (or any bully) willing money and power to AI self caricatures.

Lecture 7

(Bryson, Diamantis & Grant, *AI & Law*, 2017)

Outline

- Definitions
- Strong anthropomorphism
- Expedient anthropomorphism (Wednesday)

Summary & Future

- There is a difference between normative ethics and science.
- We can build AI that we don't owe obligations towards.
- Iff there is AI that requires obligations, arguably we shouldn't build that type of AI (Bryson 2010, 2018).
- Next lecture: Expedient Anthropomorphism and Employment (Wednesday).

Thanks (for help with legal personhood)

Tom Dale Grant

Mihailis E. Diamantis