



UNIVERSITÄT  
BAYREUTH

# Moral Status Animals and AI

**Ethics**

Matthias Brinkmann ([matthias.brinkmann@uni-bayreuth.de](mailto:matthias.brinkmann@uni-bayreuth.de))  
University of Bayreuth, Summer 2025

## 2 Today

- I. Moral Status
- II. Theories of Moral Status
- III. Degrees of Moral Status
- IV. The Moral Status of AI
- V. Parting Observations

### 3 What is Ethics?

**Ethics** investigates moral beliefs

- The aim of investigation is not empirical or descriptive: we are not interested in listing what different people or different cultures believe
- The aim of investigation is to figure out which moral beliefs are accurate, or have a higher claim to accuracy

## 4 How does one 'do' Ethics?

### In everyday moral thinking,

- we tend to think about morality often with respect to **specific cases**, but they remain unconnected
- we accept some **moral principles** and rules, but find it hard to formulate them
- ethical beliefs are part of our identity and connected to **strong emotions**

### In ethics,

- we try to build a **general theory** which applies across a range of cases
- theories are 'tested' against **hard cases**, which might involve thought experiments and edge cases
- part of the aim of building a theory is to make our concepts **more precise**
- we take a **step back** from our commitments, and consider the matter impartially

# Moral Status

---

# 6 Introduction

- Consider some moral principles
  - ‘It is wrong to insult others’
  - ‘People have a right to have their basic needs fulfilled’
  - ‘Everyone is morally equal’
- But who are ‘others’, ‘people’, ‘everyone’? The answer is provided by a theory of moral status: **what entities do morally matter?**
- Our answer has concrete implications:
  - If ChatGPT has moral status, then it is wrong to insult ChatGPT
  - If chimpanzees have moral status, then chimpanzees, even if living in the wild, have a right to have their basic needs fulfilled
  - If cats have moral status, then if you can save a cat or a human, then both choices are equally good

## 7 Moral Status: Definition

Some entity has moral status = we have **pro tanto moral reasons** to protect the entity and promote its interests **based on intrinsic properties of** that entity

- **Pro tanto moral reasons**: something that counts in favour of doing something, but can be outweighed by other factors (opposite: all-things-considered reason)
  - You have a pro tanto reason to eat healthy food, but it can be outweighed by an all-things-considered reason to eat cake at a birthday party.
- **Based on intrinsic properties**: intrinsic properties are something that an entity has by itself, not in relation to other entities
  - You have a strong reason to save your mother over a stranger, but this reason is not based on her intrinsic properties.
  - Your son's teddy bear: it would be morally wrong to let anything happen to the teddy bear, but not because of any intrinsic properties of the teddy bear
- There might be degrees of moral status (we return to this idea)

## 8 Example: Kant on Animals

- Kant thinks that **only humans** have **moral status**; thus, the categorical imperative only applies to humans (or, potentially, rational aliens)
- Nonetheless, we have **indirect duties** not to treat animals badly, e.g., because treating animals badly makes it more likely that we treat humans badly

If a man shoots his dog ... he does not fail in his duty to the dog ... but his act is inhuman and damages in himself that humanity which it is his duty to show towards mankind. If he is not to stifle his human feelings, he must practice kindness towards animals, for he who is cruel to animals becomes hard also in his dealings with men. (AA 27: 459)



# Theories of Moral Status

# Brainstorming

Why and when might entities have moral status? Think specifically of humans: why do humans have moral status?

# 11 (Some) Possible Grounds of Moral Status

1. Genetic Code/Species Membership
2. Free Will/Moral Agency
3. Cognitive Ability
4. Sentience

# 12 Human Species

**Genetic Theory.** Moral status rests on possessing human genetic code (or more generally: being human)

- Objection 1. This would entail that embryos and fetuses have moral status; corpses and severed body parts too
- Objection 2. This would exclude aliens from possessing moral status, even if they had the same abilities, skills, consciousness, etc. as us
- Objection 3. Possessing a certain type of genetic code, or belonging to a certain biological species, does not look like the right type of property to bear moral significance by itself

# 13 The Requirement of Neutrality

- Analogy
  - **Racism:** some people are less valuable merely because they belong to a certain race (where race is a quasi-biological concept)
  - **Speciesism:** some biological species are less valuable merely because of their biology
- Speciesism seems a moral error just like racism because it makes moral distinctions where there should be none
- **Requirement of Neutrality:** A good theory of moral status should not privilege humans (or any biological species) from the outset

# 14 Free Will

**Free Will Theory.** Moral status rests on possessing free will.

Kant/Descartes: humans have free will, while non-human animals do not ('animal machines')

- Objection 1. If it turned out that humans had no free will, then humans would also have no moral status
- Objection 2. How do we know that animals have no free will? How do we know that other humans have free will?

# 15 The Requirement of Observability

- **Requirement of Observability:** A good theory of moral status should rest moral status on something that we can empirically observe, at least in principle.
  - Free Will Theory seems to fail the Requirement of Observability

# 16 Cognitive Ability

**Cognitive Ability Theory.** Moral status rests on possessing cognitive abilities.

- Candidates: intelligence, rationality, language, self-awareness, etc.
- This theory fulfils both the Requirement of Neutrality and the Requirement of Observability (why?)
- We could set a **low** or a **high** threshold for cognitive abilities
  - Low-threshold cognitive ability theory: moral status rests on possessing basic cognitive abilities—e.g., match shapes, use tools, ...
  - High-threshold cognitive ability theory: moral status rests on possessing advanced cognitive abilities—e.g., possess language, think abstractly, formulate long-term plans, ...

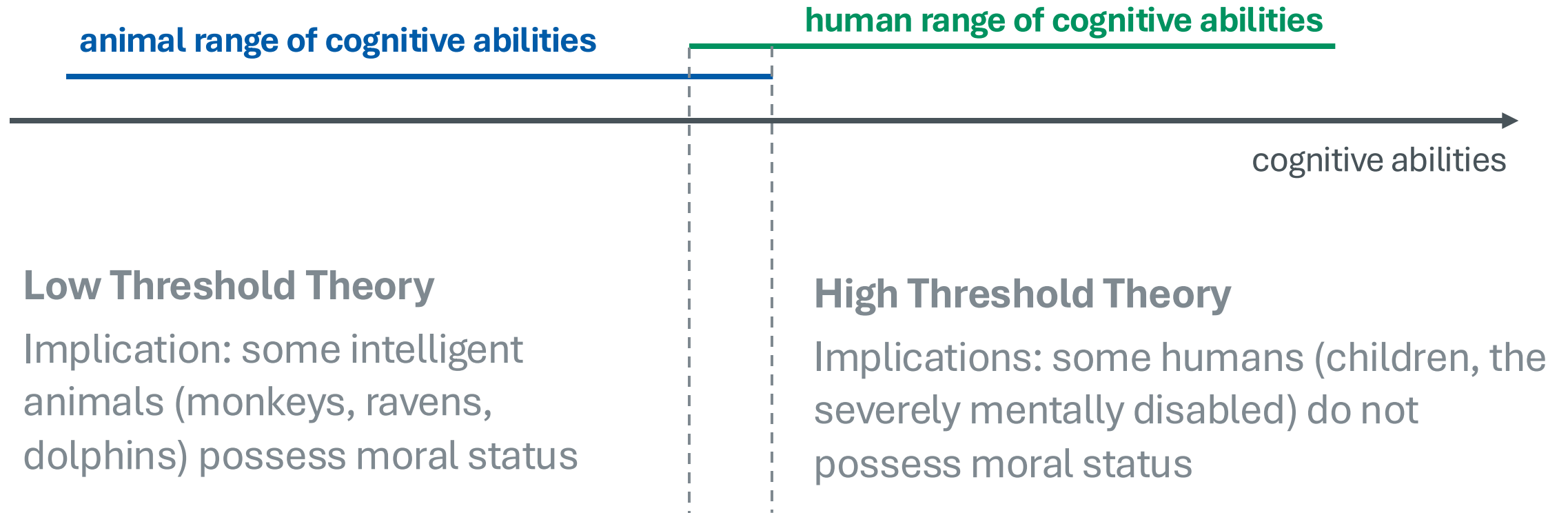


# 17 The Overlap Problem

Some national parks in the United States have a problem. Bears wander into campgrounds and break into the garbage bins. This puts both bears and people at risk. So the Park Service started installing garbage cans that were difficult to open. But it turns out it's actually quite tricky to get the design just right. Make it too complex, and people can't get them open.

Said one park ranger, **'There is considerable overlap between the intelligence of the smartest bears and the dumbest tourists.'**

# 18 The Overlap Problem



# Discussion

Which version of the Cognitive Ability Theory seems more plausible?

## 20 Capacity Theory

**Capacity Theory.** Moral status rests on possessing the potential for advanced cognitive abilities.

- This resolves some of the problems with the Overlap Problem
  - An average child has the **capacity** for advanced cognitive abilities, even though they do not currently possess those abilities
- Objection 1. But some humans do not have the potential to develop advanced cognitive abilities (e.g., severe cognitive disability due to genetic defect)
- Objection 2. Embryos possess the potential to develop advanced cognitive abilities, so they would possess moral status; but then abortion would be wrong
- Objection 3. I have the potential to become famous, but this does not mean I should currently be treated like I was famous

# 21 Species-Potential Theory

**Species-Potential Theory.** Moral status rests on belonging to a species possessing the potential for advanced cognitive abilities.

- Children and the severely cognitively impaired belong to a species possessing the potential for advanced cognitive abilities
- Objection 1. I'm dumb but all my friends are smart; but why does this mean that I should be treated like I was smart?
- Objection 2. Imagine a genetically modified cat reaches human levels of cognitive abilities; but she would have no moral status according to this theory
- Objection 3. This theory again puts moral importance on belonging to a species

## 22 Bentham: Sentience

The French have already discovered that the blackness of skin is no reason why a human being should be abandoned ... to the caprice of a tormentor. It may come one day to be recognized, that the number of legs, the villosity of the skin, or the termination of the ossacrum, are reasons equally insufficient for abandoning a sensitive being to the same fate. What else is it that should trace the insuperable line? Is it the faculty of reason, or perhaps, the faculty for discourse? ... the question is not, Can they *reason*? nor, Can they *talk*? but, Can they *suffer*? (Jeremy Bentham)

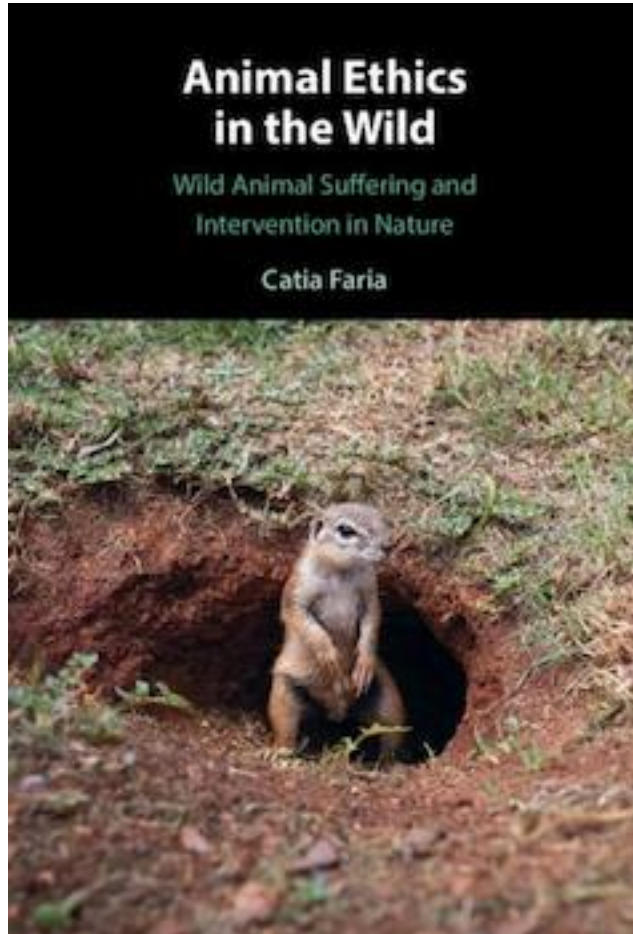
# 23 Sentience Theory

**Sentience Theory.** Moral status rests on being sentient.

**Ability-to-Suffer Theory.** Moral status rests on being able to suffer.

- Sentience = the capacity to experience feelings, emotions, and sensations
- Mammals and birds experience pain, as do other animals; fetuses probably develop sentience in the third trimester
- This theory (partially?) fulfils the Requirement of Observability
  - Example: Anxiety in humans is caused by benzodiazepine receptors; we know that many vertebrate species have such receptors, and that they are active when these animals appear to act anxiously
  - Example: most insects do not appear to exhibit pain-induced behaviour (e.g., unchanged behaviour after limb has been severed)

## 24 Objections to Sentience Theory



**Sentience Theory.** Moral status rests on being sentient.

- Objection 1. According to this theory, we also have reason to help wild animals and intervene in nature
- Objection 2. Sentience is too low a bar
  - Imagine a species which only experiences the moment, and has no way to conceptualise what happens to it
- Objection 3. Sentience is too high a bar (examples later)
- Objection 4. Because both mice and humans are sentient, both have equal moral status. Thus, we have stronger reason to save two mice rather than one human

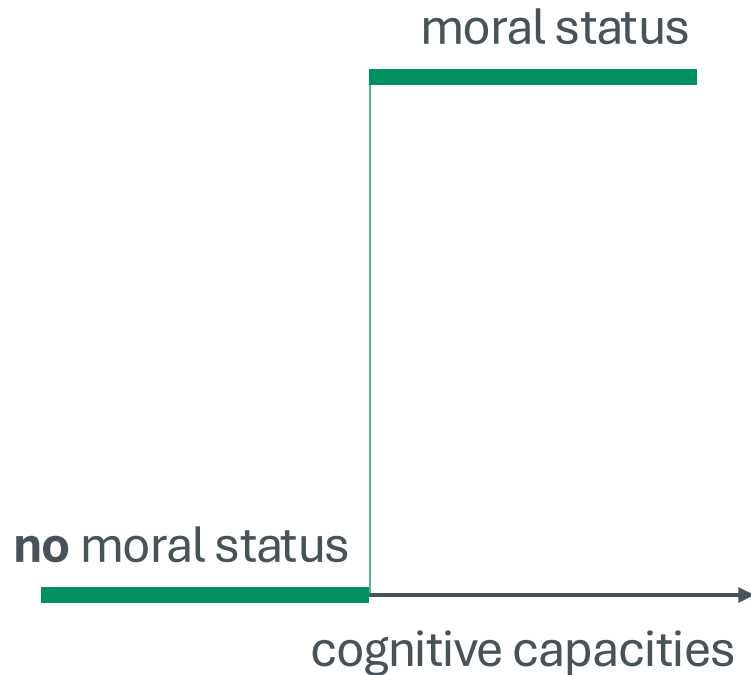


# Discussion

What speaks in favour of sentience theory?  
How might we be able to fix its problems?

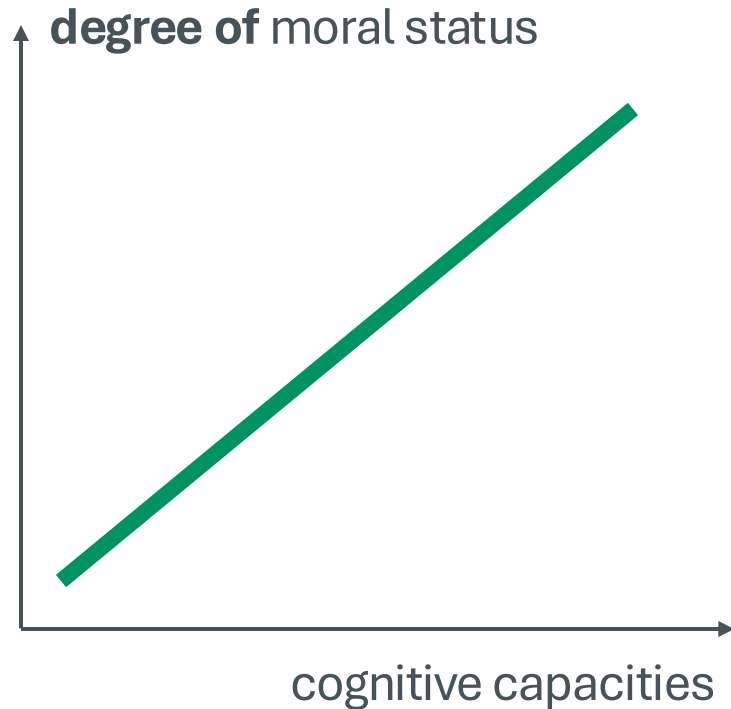
# Degrees of Moral Status

## 27 Binary Moral Status



- A property is **binary** if the property either exists or does not ('is alive', 'is a mammal')
- A property is **scalar** if the property exists to various degrees ('is 181 cm tall')
- The theories considered so far specify moral status as something binary
- But given that the grounds of moral status (e.g., cognitive capacities) are scalar, maybe moral status is also scalar
- Perhaps some of the problems with existing theories can be fixed this way

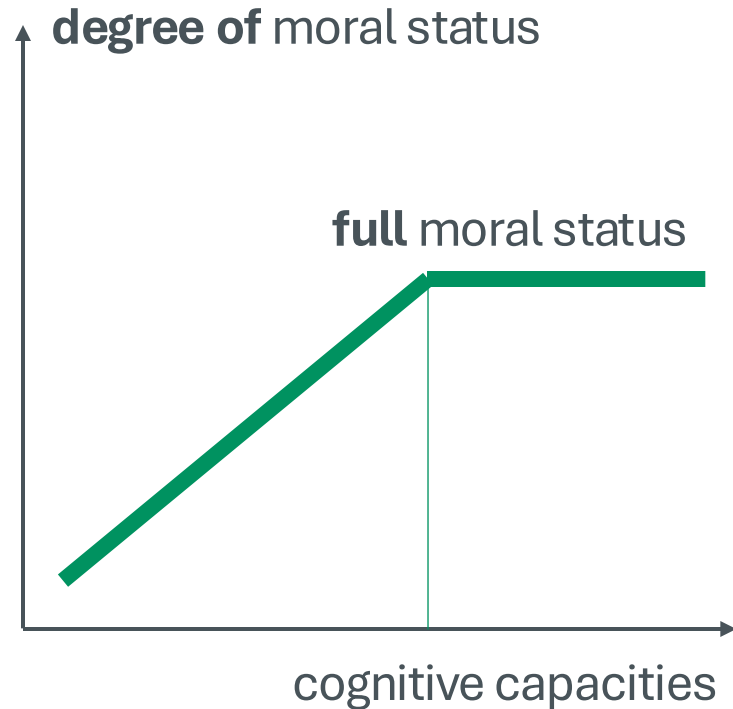
## 28 Scalar Moral Status



**Scalar Cognitive Capacity Theory:** the more cognitive capacity an entity possesses, the higher its degree of moral status

- Objection 1. On this theory, some non-human animals (e.g., monkeys) have higher moral status than some humans (e.g., children)
- Objection 2. On this theory, some humans have higher moral status than other humans
- Objection 3. What does it mean that something has higher degrees of moral status?

## 29 Scalar Moral Status with Upper Bound



**Scalar Cognitive Capacity Theory (II):** the more cognitive capacity an entity possesses, the higher its degree of moral status, until a maximum amount of moral status is reached

- Objection 1. We have problems with specifying the threshold again
- Objection 2. Why should beings with much greater capacities (e.g., angels, superintelligent aliens) have the same moral status as us?
  - Imagine: the cognitive distance between aliens and us is as big as the distance between us and animals

# Discussion

Which theory of moral status seems the most appealing? Which avoids the most philosophical problems?

# The Moral Status of **AI**

## 32 Why does it matter?

- Is it morally wrong to verbally abuse a chatbot?
- Is it morally wrong to kill and harm NPCs in computer games?
- Would it be murder to pull the plug on a supercomputer?
- Should we not kick a robodog?



## 33 Lt Cmdr Data

A fictional android in *Star Trek: The Next Generation*

- Artificially created and programmed
- Super-human intelligence
- Independent, self-chosen life aims (specifically, to become more human)
- Interacts socially
- Moral agency
- Incapable of emotions
- Incapable of feeling pain

# 34 Does Data Possess Moral Status?

(TNG 2:09,  
“The Measure  
of a Man”)

## 35 Reconstructing the Argument

- Proposed Theory: Moral Status rests on Sentience (intelligence, self-awareness, consciousness) (careful: 'sentience' here is used differently!)
- Proposed Claim: Data does not have moral status according to this theory
- Counterargument: Data fulfils self-awareness and intelligence criteria; consciousness does not satisfy the Observability Requirement, and so cannot be required

## 36 A Continuum of AIs?

- Robot vacuum cleaners
- NPCs (non-player characters) in computer games
- Chess engine (e.g. Stockfish)
- Personal assistant (e.g. Alexa, Siri)
- Autonomous vehicles (e.g. Waymo)
- LLMs (e.g. ChatGPT)
- Dolores (from *Westworld*)
- Data (from *Star Trek*)

# 37 Challenges to Existing Theories

## 1. Genetic Code/Species Membership

⇒ The example of AI makes it even clearer that genetic code or biological membership in a species isn't a good ground for moral status

## 2. Free Will/Moral Agency

⇒ Free Will. One might think that (some) AIs are 'deterministic'; but one might ask whether humans are simply not also machines the source code of which we do not know

⇒ Moral Agency. The example of AIs probably requires us to rethink what moral agency means

## 3. Cognitive Ability

⇒ AIs possess many cognitive abilities, and are better than humans at certain tasks. So they would clearly qualify for moral status. But we might also ask whether mere cognitive ability is enough, and which cognitive abilities matter.

## 4. Sentience

⇒ In the narrow sense of feeling emotions or pain, many AIs are not sentient. But we might not ask whether this is not too narrow a picture of sentience.

# Parting Observations

# Key Lessons

- Some entity has moral status = we have pro tanto moral reasons to protect the entity and promote its interests based on intrinsic properties of that entity
- Moral status matters because it determines who is covered by which moral principles
- Three main families of theories: free will/agency; cognitive ability; sentience
- Each theory is subject to various objections
- The example of AI will require further modifications of our thinking about moral status

# 40 Tutorials & Next Week

- Text for Tutorials: Peter Singer, *Practical Ethics*, chapter on animals
- Secondary Text: Walter Sinnott-Armstrong and Vincent Conitzer, 'How Much Moral Status Could Artificial Intelligence Ever Achieve?'
- **No lecture** next week (Wittgenstein Lectures)
- **In two weeks: collective obligations**
  - What should I morally do if I do not make a difference?
  - What should I morally do if others do not do their fair share?
  - When am I required to coordinate myself with others?
  - (Think of these questions in the context of climate change...)



## 41 Extra: Which Beings Are Sentient?

There is overwhelming evidence that mammals and birds are sentient creatures. The evidence that reptiles are sentient is also quite strong; the evidence for amphibians, more intermediate. Meanwhile, the evidence that bony fishes and cephalopods (octopi, squid, and cuttlefish) are sentient is fairly strong. As for cartilaginous fishes such as sharks and stingrays, jawless fish such as lampreys and hagfish, and some invertebrates other than cephalopods such as crustaceans and insects, the evidence is mixed. For most other invertebrates (e.g., worms, jellyfish, urchins, sponges), the evidence suggests that they are probably not sentient. As for plants, although there is evidence that they can process information from their immediate environment and use that information to move their parts adaptively, we doubt there is any evidence that they actually *feel* anything.

(DeGrazia/Millum, *A Theory of Bioethics*, ch. 7)