Robots & Ethical Theory: Utilitarianism

COGS300.002

Peter Danielson

Jan 20 2015
Ethical Theory

- Robot Ethical Theory (Applied next 2 lec)
  - How does existence of new artificial moral agents and patients fit into ethical theory?

- Utilitarianism:
  - Values
  - Values and decisions
  - Constraints
    - Projects (agents)
    - Rights (patients)
Moralized Decision Theory: Right is Maximizing

• Value
  – Pleasure (Bentham)
  – Happiness (J.S. Mill)
    • Happy pig vs. dissatisfied Socrates
  – Ideal goods (G.E. Moore)
  – Preferences (modern economics)
  – Primary Goods (J. Rawls)
  – Quality Adjusted Life Years (medical cost-benefit)
Vs. Yuck and Cute

"Utilitarians ... fought for ... reason against mere tradition, dogmatism, and vested interests" (624)

2. Artificial Meat

Scientists have developed a method of growing animal protein in a laboratory. Using cloned stem cells from cows, pigs and chickens, they have managed to grow edible meat in petri dishes. The meat produced is indistinguishable from that of live animals – it tastes, looks and feels exactly the same, and has the same nutritional content.

A large food distribution company is now using this method to mass-produce artificial meat. Unlike traditional farming, no animals need to be bred, housed or killed in order to produce this meat.

Steaks, burgers and sausages made from this artificial meat are now being sold at your local supermarket and served in your local restaurant. This artificial meat costs the same as conventional meat, and has met the same governmental health requirements. It is as safe for human consumption as any other food product on the market.

Would you eat this artificial meat?

A) Strong Yes B) Weak Yes C) Neutral D) Weak No E) Strong No
Gau’s Plan

1. Del’s objection: calculating utility
   – Cuteness, QALY’s & rescue

2. Williams’ objection: personal projects

3. Rawls’ objection: rights for us;
   – Nozick: utilitarianism for (robots and) animals

• Conclusion: design constraints on robots
Easy Case

A Rescue bot can only save one group of innocent humans in an earthquake;
A It should go East and save 1.
B It should go West and save 5.
Q1

Which of the following would a utilitarian, such as VIKI, not find morally permissible?

a) A person donating their own money, originally saved up for an expensive vacation, to the local food bank.

b) The government refusing to release a medication that has cured patients to the public because of a few minor side effects.

c) A nurse that donates her dead patient's organs without the patient's consent.

d) A lifeguard saving the lives of two drowning people on one side of the pool instead of one drowning person on the other side.

e) None of the above

Tisha
Rate Quiz Question 1

A. Excellent
B. Very Good
C. Good
D. Acceptable
E. Poor
VIKI, according to the article, chose to save Del Spooner (Will Smith) instead of a little girl as he had a higher chance to be successfully rescued. Is this decision utilitarian?

A. Yes, because Del was more likely to be saved and thus had a higher ethical utility.
B. Yes, Del is more intellectually mature and hence capable of a greater degree of suffering and well being, rendering himself with a greater ethical utility.
C. No, consideration for mere rescue probabilities isn't sufficient to derive a utilitarian conclusion from.
D. No, even with a smaller probability the little girl has a greater ethical utility as she is likely to have a much longer lifespan after the rescue than Del.
E. Only Jesus knows.

Daehan
Rate Quiz Question 2

A. Excellent
B. Very Good
C. Good
D. Acceptable
E. Poor
Discussion

Should the robot save:

A. Del (with a .45 chance of success) or
B. the little girl (with an .11 chance of success)
Choice under uncertainty

• Expected Utility: Value $\times$ probability
  – $1 \times .11 < 1 \times .45$

• Exclude: Cuteness

• Include: Age
  – Girl’s QALYs $= \sim 82 - 12 = 60 \times \text{qual} = 1$
  – Del’s QUALYs $= \sim 80 - 40 = 40 \times \text{qual} = .5$

• Exclude Will’s
  – Responsibility?
  – Choice?
QALYs

- Life (value)
- Years (quantitative)
- Quality (qualitative)
- Probability
  - [http://www.scottishmedicines.org.uk/About_SMC/Policy_Statements/A_Guide_to_Quality_Adjusted_Life_Years](http://www.scottishmedicines.org.uk/About_SMC/Policy_Statements/A_Guide_to_Quality_Adjusted_Life_Years)
  - E.g which expensive medicines should a public health care system provide?
Grau refers to a scenario in which there was a drowning girl and says while the human would unreflectively but correctly "save the girl", the robot engages in calculations and deliberations that exhibit ____________ (as phrased by Bernard Williams).

A) "Fast thinking"
B) "One thought too many"
C) "Accurate strategizing"
D) None of the above

Briana
Rate Quiz Question 3

A. Excellent
B. Very Good
C. Good
D. Acceptable
E. Poor
Personal Projects vs. Demands of Morality

- Williams’ “one thought too many”
  - What is that thought?
  - Role (wife/husband) overrides impersonal morality
    - What about “pet owner” (Santa Barbara)

- W/o personal projects, robots can be held to a higher standard
  - (what about ownership?)
The integrity objection to utilitarianism states that

A) utilitarianism has difficulty recognizing the type of morality that focuses on the person acting, rather than the recipients of the person's actions
B) utilitarianism requires such a great moral sacrifice that the agent no longer really has a life at all
C) if a robot has been given a sense of human morality, that it would be immoral for us to require it to act in a utilitarian way
D) we ought to create a sense of human morality (exemplified by sentience, autonomy, a sense of self) in robots if we have the means to do so

Michael W
Rate Quiz Question 4

A. Excellent
B. Very Good
C. Good
D. Acceptable
E. Poor
Rights

• Constrain Value Maximizing
  – Would be *best* but free speech, due process, ownership...

• Rawls’ contract constraint on utilitarianism
  – Separate persons
  – Restrict sacrifice of one for other(s)
    • Absolutely for human rights
    • unless all (worst off) benefits for economics
  – But would we choose to restrict organ transfers?
    • E.g avoidable kidney deaths due to opt in consent
Applying Rights

• But our results on Human Trolley1/Switch:

<table>
<thead>
<tr>
<th>Option</th>
<th>Count</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>24</td>
<td>80%</td>
</tr>
<tr>
<td>B</td>
<td>5</td>
<td>17%</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>3%</td>
</tr>
</tbody>
</table>

Kill 1       Let 5 Die       ???
Discussion

I thought that the results from this multi step question would be cool to see.

Imagine the trolley problem from class. The trolley is headed towards the track with five people but you have the option to pull the lever so that the trolley changes tracks and will strike only the one person. What do you do?

A. Let trolley strike the five people
B. Pull the lever and let the trolley strike the one person
Now imagine the same problem but this time the five humans are replaced with five robots (who do have a sense of individual identity). However, the one person on the other track is still human. What do you do?

A. Let the trolley strike the five robots
B. Pull the lever and let the trolley strike the one human

Jas
Rate Quiz Discussion Q

A. Excellent
B. Very Good
C. Good
D. Acceptable
E. Poor
Robot Agent 1

Trolley with with robot driver should:

A. Let the trolley strike the five humans
B. Pull the lever and let the trolley strike the one human
Robot Agent 2

Trolley with robot driver should:

A. Let the trolley strike the five robots
B. Pull the lever and let the trolley strike the one human
Q5

According to Grau, in which situation should a robot use utilitarian reasoning?

A) A medical robot deciding who to treat first.
B) A sentient robot deciding whether to sacrifice itself to save humans.
C) A robot removing a key part and giving it to [save] another robot.
D) A sentient robot with a sense of individuality playing a game with humans.

Michael Z
Rate Quiz Question 5

A. Excellent
B. Very Good
C. Good
D. Acceptable
E. Poor
Q6

Of the options below, which is/are NOT suggested by Grau as relevant for a robot's possession of an "individual identity".

A) Self-awareness and self-governance
B) The ability to recognize and respond to reasons
C) The means to apply utilitarian rationality
D) The capacity for free and responsible choice
E) B and C

Eric
Rate Quiz Question 6

A. Excellent
B. Very Good
C. Good
D. Acceptable
E. Poor
Nozick’s Two Status Moral Sketch

Rights for Humans
• “Humans may not be used or sacrificed for the sake of others”
• Recall Asimov’s laws

Utilitarianism for animals:
• “Animals may be sacrificed for the benefit of other people or animals only if the total benefits are greater than the ... loss”
• And, Gau adds: some robots
Gau’s Arguments/Conclusions

1. Williams’ objection: personal projects & robot agents
   – But robots have no life projects to exempt them from higher moral standards
   – Therefore we ought not make robots into persons

2. Rawls rights for us
   – But utilitarian robots apply should be constrained by rights

3. Nozick: utilitarianism for animals & robot patients
   – Except in case of robots and animals where unrestricted utilitarianism is appropriate
   – Again: a reason to not develop fully conscious robots
The Grau article critiques the use of utilitarianism for interactions with humans, and says that a utilitarian robot "would not be a suitable ethical advisor for humans when considering acts that affect other humans."

But while few people are pushing for society to become completely utilitarian, utilitarian reasoning is used all the time; Couldn't the utilitarian robot's analysis be useful as one thing for a human to consider when making an ethical decision in regard to an action affecting another human?

Paul
Rate Discussion Ques 2

A. Excellent
B. Very Good
C. Good
D. Acceptable
E. Poor
Discussion Question 3

What moral obligations, if any, do humans owe to robots or potential AI designs?
And who is responsible for enforcing these or any other robot design obligations/limitations (architects, designers, owners, governments)?

Wesley
Rate Discussion Ques 3

A. Excellent
B. Very Good
C. Good
D. Acceptable
E. Poor
Rate the Reading Assignment: Grau

A. Excellent
B. Very Good
C. Good
D. Acceptable
E. Poor
Next

• Ethics of Lethal Military Robots

• Robot Ethics Survey recommended
References

