

Machine Supererogation & Deontic Bias

Jonathan Pengelly

CPG Annual Conference 2019

Governing the Future: Digitalization, Artificial Intelligence, Dataism

October 19, 2019

My Research Focus

Exploring the possibility of "widening the circle of moral participation" to establish **some form** of moral community with both human and artificial moral agents.

Outline

Outline

1. Is machine supererogation possible?

Outline

1. Is machine supererogation possible?
2. A deontic bias exists within machine ethics

Supererogation

A supererogatory act is ...

A supererogatory act is ...

- a morally admirable act that in some way goes "beyond the call of duty."

A supererogatory act is ...

- a morally admirable act that in some way goes "beyond the call of duty."
- an action that is praiseworthy if performed, but not blameworthy if omitted.

Paradigm Examples

- Heroism and Saintliness

Paradigm Examples

- Heroism and Saintliness
- Beneficence and Charity

Paradigm Examples

- Heroism and Saintliness
- Beneficence and Charity
- Volunteering

Paradigm Examples

- Heroism and Saintliness
- Beneficence and Charity
- Volunteering
- Forgiveness

The Paradox of Supererogation

The Paradox of Supererogation

- How can the moral good not be required as a duty?

The Paradox of Supererogation

- How can the moral good not be required as a duty?
- Anti-supererogationism vs Pro-supererogationism

*"The problem of supererogation permits classical and current theories of morality to be **explored from an unusual angle**, and such an exploration may underline both some merits and some flaws and inconsistencies in those theories."*

Heyd, *Supererogation: Its Status in Ethical Theory*

Machine Supererogation

Licentia

- "Allowing" an agent to exercise their power of moral choice

Licentia

- "Allowing" an agent to exercise their power of moral choice
- *"More virtue in doing what is not required of us than in doing what is really no more than our duty"* Heyd, *Supererogation*

Licentia

- "Allowing" an agent to exercise their power of moral choice
- *"More virtue in doing what is not required of us than in doing what is really no more than our duty"* Heyd, *Supererogation*
- *"I have done "my duty" now my time, my choices are all mine"* Baron, *Kantian Ethics and Supererogation*

The Two Faces of Morality

- Basic duties vs "the higher flights of morality"

The Two Faces of Morality

- Basic duties vs "the higher flights of morality"
- "*There seems to be a limit to how much morality we can stand*"
Wolf, *Moral Saints*

The Personal Dimension of Morality

- *"There is a lightness and ease about the performance of such acts"*

The Personal Dimension of Morality

- *"There is a lightness and ease about the performance of such acts"*
- *"Love is personal and particular and makes demands that are not quantifiable in standard moral theories."*

Drummond-Young, Is Supererogation more than just costly sacrifice?

Deontic Bias in Machine Ethics

"But deontic logic is not the logic of morality; it is instead the logic of rights and duties, the logic of right conduct; and that which is neither required nor forbidden is therefore shown to be indifferent only with respect to rights and duties; not ... with respect to morality itself."

Schumaker, *Deontic Morality and*

the Problem of Supererogation

We use a rich moral language representing moral concepts incompatible with a purely deontic interpretation of morality. However ...

We use a rich moral language representing moral concepts incompatible with a purely deontic interpretation of morality. However ...

we frame many debates within machine ethics in a deontic way.

*"They act as **linguistic Trojan horses** that smuggle in a rich interconnected web of human concepts that are not part of a computer system or how it operates. Once the reader has accepted a seemingly innocent Trojan term ... it opens the gate to other meanings associated with the natural language use of the term"*

Sharkey, The evitability of autonomous robot warfare

Machine Moral Responsibility

"There is nothing strange about the idea that morality asks more of us than we should demand of each other."

Ferry, *Beyond Obligation: Reasons and Supererogation*

Recommendations

- We have to be more careful with our moral language.

Recommendations

- We have to be more careful with our moral language.
- Explore the richness of human morality.

Summary

- Is machine supererogation possible?
 - Licentia
 - The Two Faces of Morality
 - The Personal Dimension of Morality
- A deontic bias exists within machine ethics

Thanks for your attention!

Any questions?

Email: Jonathan.Pengelly@vuw.ac.nz