



Matthias Uhl (matthias.uhl@thi.de)

Russian-Bavarian Online Workshop "Artificial Intelligence and Law"

June 17, 2021

7

Agenda

- 1. Behavioral Ethics
- 2. Why is this important for Human-Machine Interaction?
- 3. Shaping our moral conduct or the Ethics of Things



1. Behavioral Ethics



1. Behavioral Ethics?

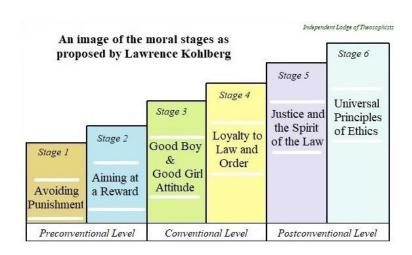
Traditional ethics ascribes central role to deliberation

Taking normative approach, focus is on how we should act when resolving ethical dilemmas

Usually, it is assumed that once people have understood what is right and what is wrong, they act accordingly

Centrality of deliberation for ethics is exemplified by Kohlberg's (1969) stages of moral development: morality develops through different stages and then becomes stable trait

Most of us dramatically underestimate how strongly our behavior is prone to incentives and other situational factors





1. Behavioral Ethics?

Most of us underestimate how much our behavior depends on situational factors

Behavioral ethics examines how people actually behave and assumes that they often act in contradiction to their own ethical beliefs

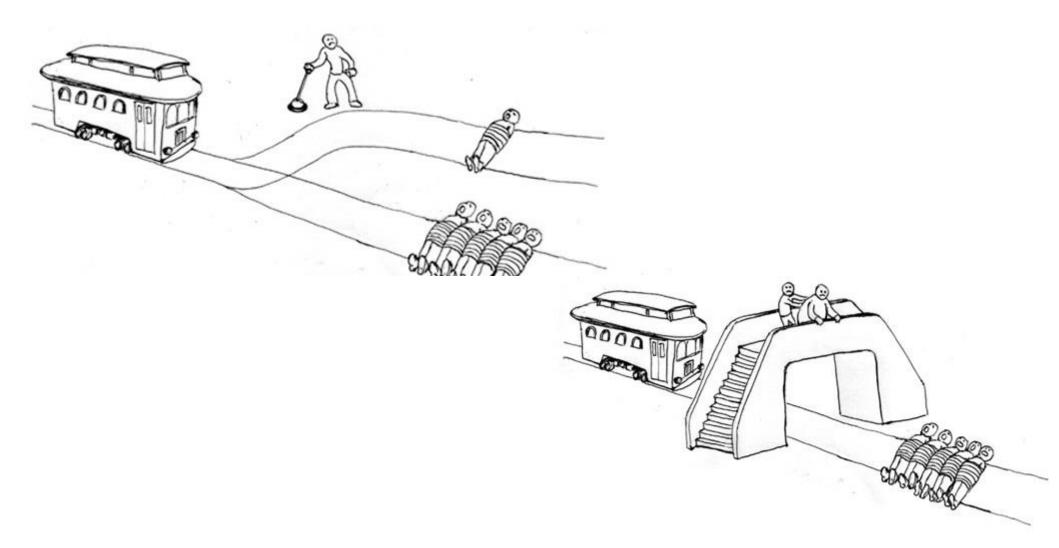
Traditional ethics lacks an understanding of the non-intentional cognitive factors influencing immoral behavior

It is important to study these cognitive influences scientifically because they are not idiosyncratic "noise" in the data, but predictable patterns that can be addressed



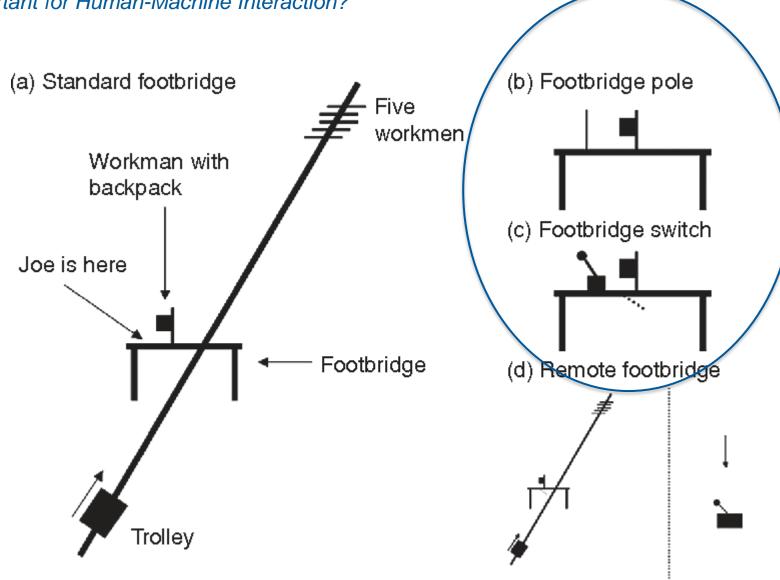
2. Why is this important for Human-Machine Interaction?

Why is this important for Human-Machine Interaction?





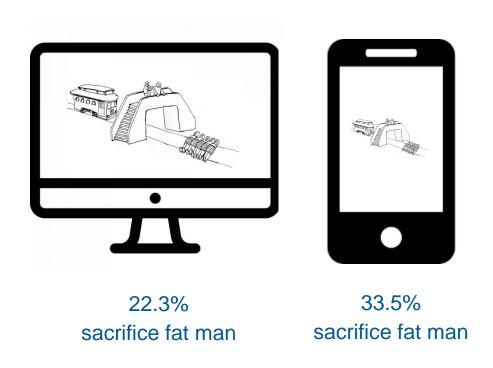
2. Why is this important for Human-Machine Interaction?



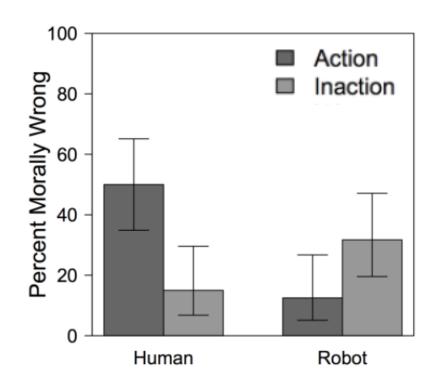
Greene (2013)

4

2. Why is this important for Human-Machine Interaction?



Barque-Duran et al. (2017)



Voiklis et al. (2016)



3. Shaping our moral conduct or the Ethics of Things



3. Shaping our moral conduct or the Ethics of Things

The neutrality thesis of technology is falsified by empirical evidence

We have a responsibility for an ethically aligned design

Setting defaults is one way to influence behavior

One may use this to nudge people in the "right direction"

One may also decide intentionally to abstain from doing so, for instance, by randomizing defaults

In any case, designers of technology have to accept their responsibility

