Lect. #9: Ethical Theories 1

dilbert.com/strips/comic/2006-07-04

Agenda for Today

1. Autonomous vehicles: Issues and Ethics

2. Ethics vs Morality

3. Moral Systems

4. Ethical Theories: Social Utility and Consequence-Based Theories

Self-Driving Cars (AKA Driverless Cars, Autonomous Vehicles): When?

California regulators approve state's 1st robotic taxi fleet

By MICHAEL LIEDTKE June 2, 2022

Home 〉 News 〉 Oxford researchers develop new AI to enable autonomous vehicles to adapt to challenging weather conditions

Oxford researchers develop new AI to enable autonomous vehicles to adapt to challenging weather conditions

PUBLISHED
8 SEP 2022

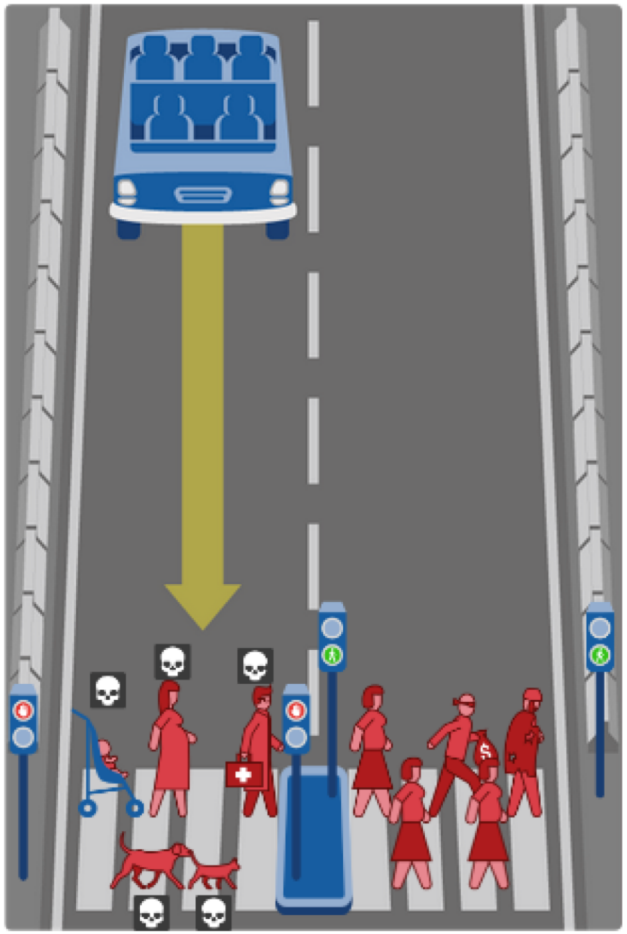
SHARE THIS

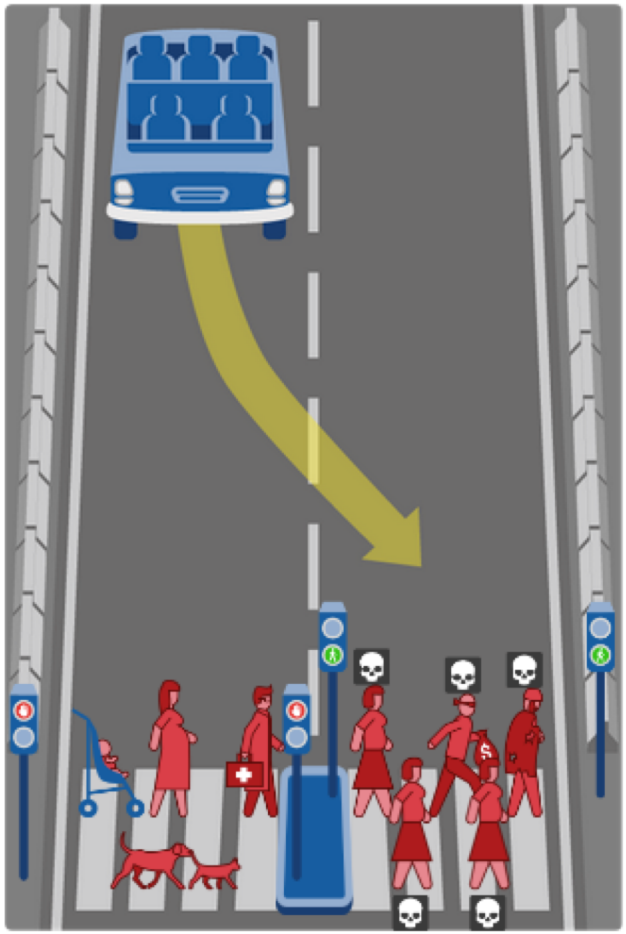
Driverless Cars Raise Ethical Issues for Engineers

freeths.co.uk

- Algorithms determine the behavior of driverless cars
- Suppose it's your responsibility to specify what the algorithms would do
- How will you decide what to specify, and what will you specify?

Your Specification Would Have to Handle the Following Scenario





moralmachine.mit.edu

Poll

Would your algorithm run into those on the left or those on the right?

left: a baby, a male doctor, a pregnant woman, a cat, and a dog, all ignoring the traffic signal

right: three large women, a criminal, and a homeless person, all obeying the traffic signal

A. The three/four humans and two animals on the left

B. The five humans on the right

Breakout Discussion (3 min)

- Discuss the two choices. What did you choose and why?
- Choose someone to present your discussion results including the number of your group who chose left and the number who chose right and the main reasons for these choices

Poll

Would your algorithm run into those on the left or those on the right?

left: a baby, a male doctor, a pregnant woman, a cat, and a dog, all ignoring the traffic signal

right: three large women, a criminal, and a homeless person, all obeying the traffic signal

A. The three/four humans and two animals on the left

B. The five humans on the right

Ethics vs Morality

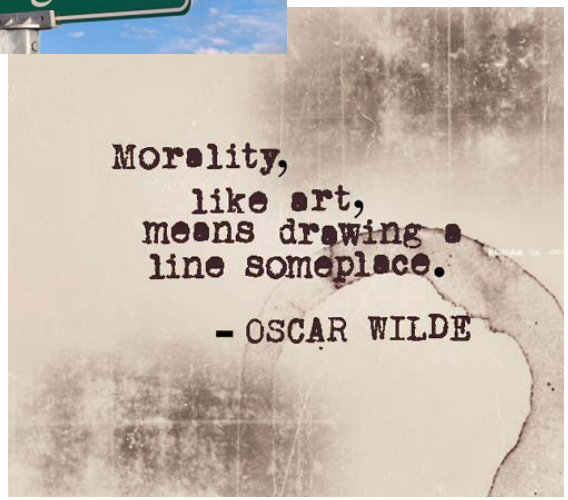
- What is ethics? Ethics is the study of morality.
- What is morality? **Morality is a system of rules** (mores) that distinguish between right and wrong
- Ethics comes from the Greek word *ethos* which means nature or disposition
- Morality comes from the Latin word *moralis* which itself comes from the Latin word *mos* (pl. *mores*), which means custom

wiringthebrain.com/2011/06/where-do-morals-come-from.html

Morality Is a System of Rules = Moral System

- At micro-level:
  - Personal directives, e.g.,
    - Don’t steal.
    - Don’t cheat.
    - Don’t harm others.
- At macro-level:
  - Social policies and laws, e.g.,
    - Privacy policies (Facebook)
    - Copyright laws
    - Professional codes of conduct
    - Government regulations

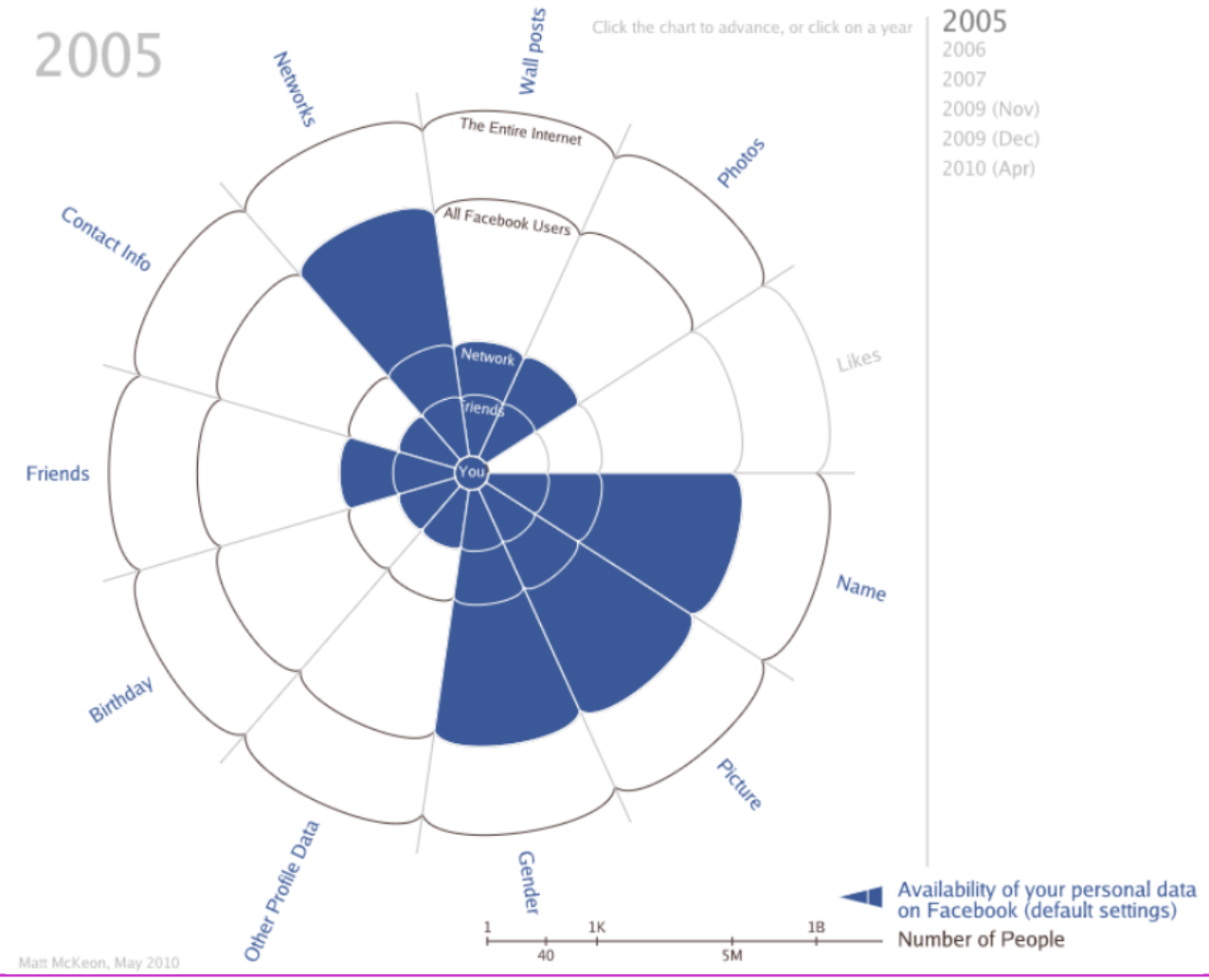
youtube.com



iheartinspiration.com

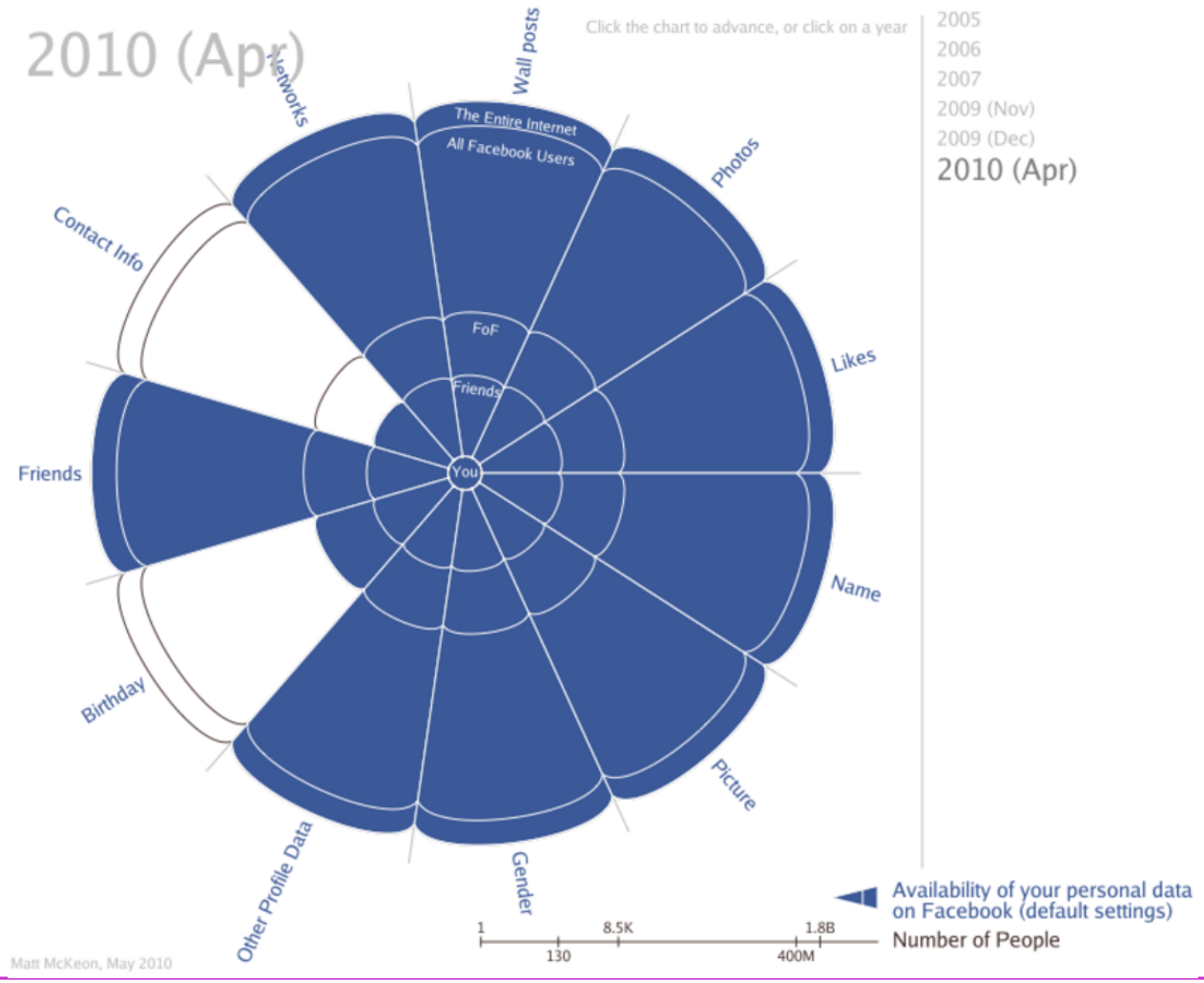
Default Privacy Settings in Facebook: 2005

mattmckeon.com/facebook-privacy



Default Privacy Settings in Facebook: 2010

mattmckeon.com/facebook-privacy



Basic Components of a Moral System

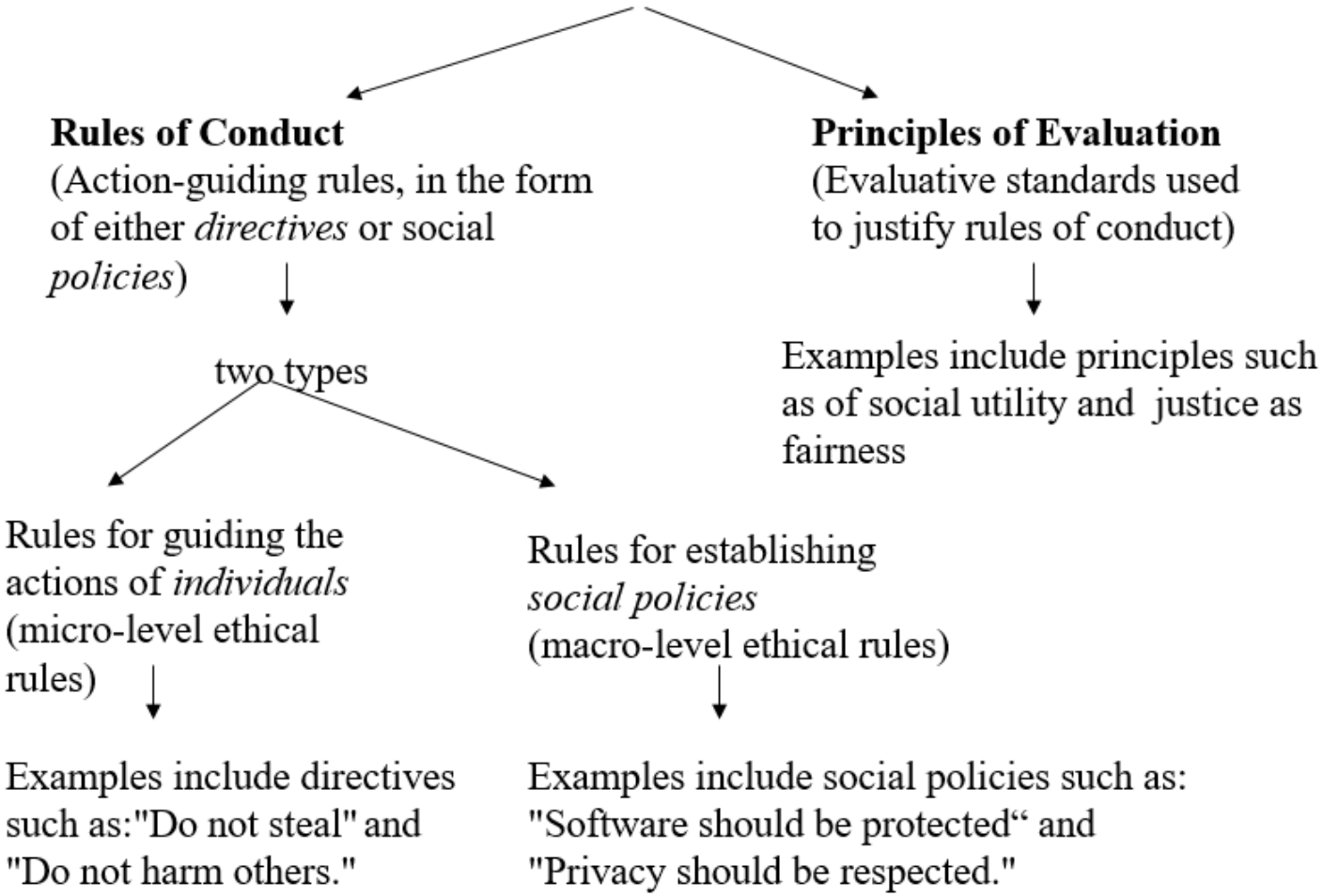


Fig. 2.1, Tavani

Moral Systems Are Evaluated Using Principles

- Principle of social utility
  - Jeremy Bentham (1748-1832)
  - Promoting the greatest good for the greatest number
  - Use to evaluate both micro- and macro-level rules of conduct, e.g.,
    - Proprietary software shouldn't be copied without permission

learnliberty.org

Poll

Using the principle of social utility, now what would your algorithm do—run into those on the left or those on the right?

left: a baby, a male doctor, a pregnant woman, a cat, and a dog, all ignoring the traffic signal

right: three large women, a criminal, and a homeless person, all obeying the traffic signal

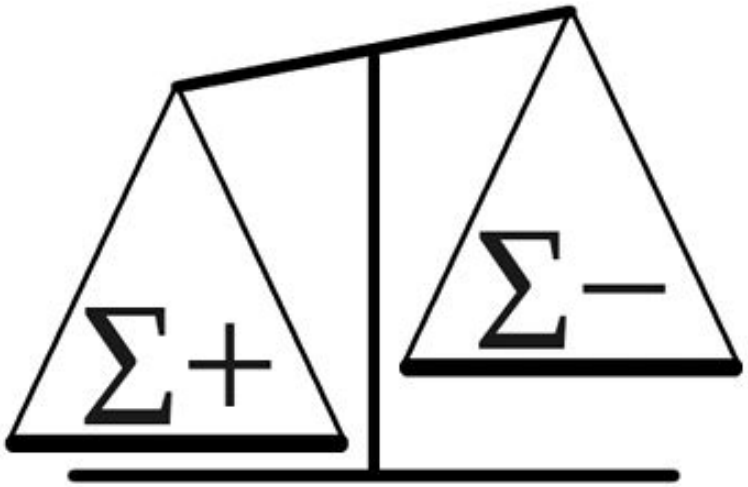
A. The three/four humans and two animals on the left
B. The five humans on the right

Brief Overview of Four Ethical Theories

- Consequence-based: Based on the principle of social utility; act to yield the greatest good for the greatest number
- Duty-based: Based on the principle of duty or obligations human have to one another and never on the consequences of human behavior
- Contract-based: Based on the idea that humans should act to satisfy their own natural desires, but because humans are rational, this will result in a system that will protect them from being harmed by others
- Character-based: Based on the idea that humans with strong moral character will do the right thing

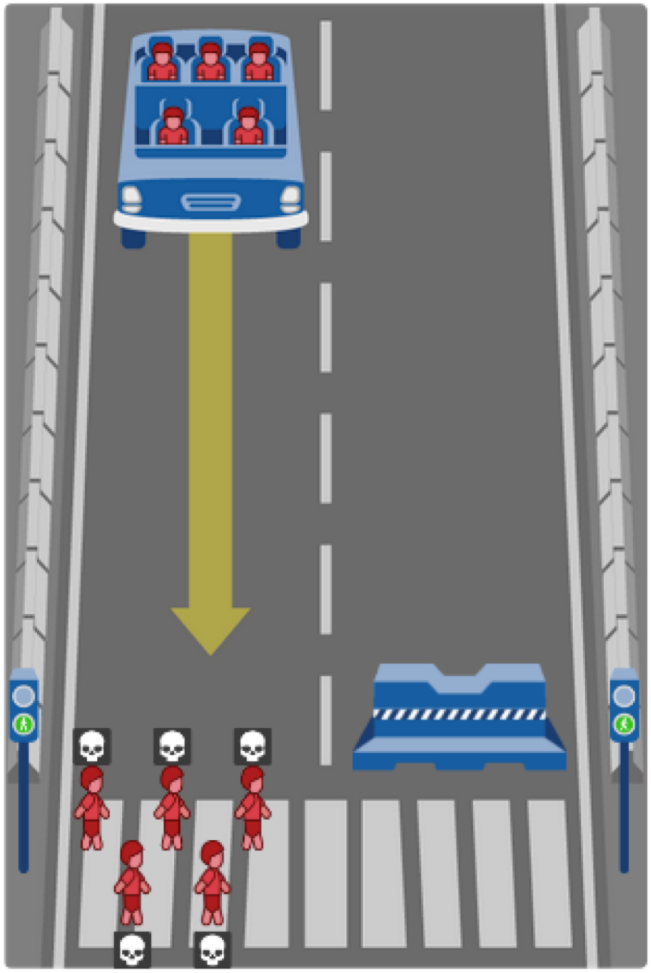
Act Utilitarianism
Jeremy Bentham (1748-1832)

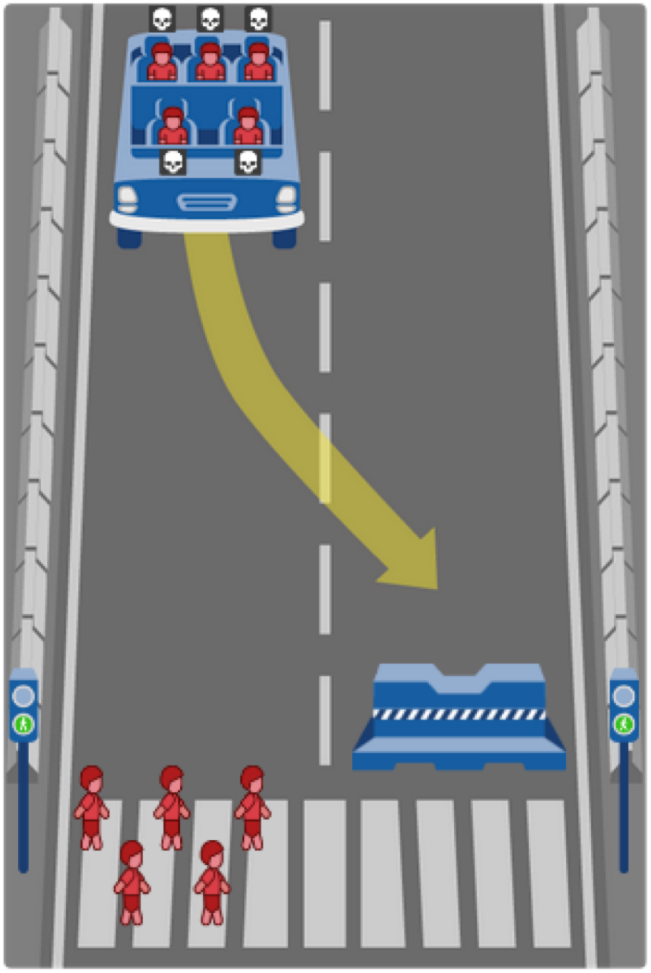
- An act is ethically acceptable if its consequences result in the greatest good for the greatest number of people
- Construct a stakeholder table with stakeholders in the first column, relative weight of each stakeholder in the second [1:5], weight of positive or negative outcome in the third [-5:5], and reasoning in the fourth. Evaluate by taking sum of products of weights.



powerlisting.wikia.com

Using Act Utilitarianism to Evaluate Autonomous Vehicle Scenario





moralmachine.mit.edu

Act Utilitarianism: Stakeholder Table

Scenario: Assume Passengers Are Killed

Stakeholder	Weight	Net Utility	Reasoning
Passengers in Vehicle			
Pedestrians			
Car Designer/ Manufacturer			
City government			

Who's involved in this scenario?

Utilitarian analysis:

Act Utilitarianism: Stakeholder Table

Stakeholder	Weight	Net Utility	Reasoning
Passengers in Vehicle	1		
Pedestrians	1		
Car Designer/ Manufacturer	3		
City government	3		

What's the relative weight of each stakeholder involved [1-5], where 1 is a single person and 5 is larger entity, e.g., general public?

Utilitarian analysis:

Act Utilitarianism: Stakeholder Table

Stakeholder	Weight	Net Utility	Reasoning
Passengers in Vehicle	1	-5	
Pedestrians	1	5	
Car Designer/ Manufacturer	3	0	
City government	3	0	

What's the weight of the positive or negative outcome of the scenario [-5:5]?

Utilitarian analysis:

Act Utilitarianism: Stakeholder Table

Stakeholder	Weight	Net Utility	Reasoning
Passengers in Vehicle	1	-5	All passengers die, but perhaps there's some utility because they sacrificed so others could be saved
P			All pedestrians live, but perhaps there's some negative utility knowing that those in the car died
C M			Difficult to assess how algorithm's choice will impact company sales, but it will undoubtedly look bad that car was involved in accident
City government	3	0	City government may receive backlash from public; if a softer barrier had been used, it may have saved lives

For each stakeholder, describe your reasoning for assigning the weights.

Utilitarian analysis:

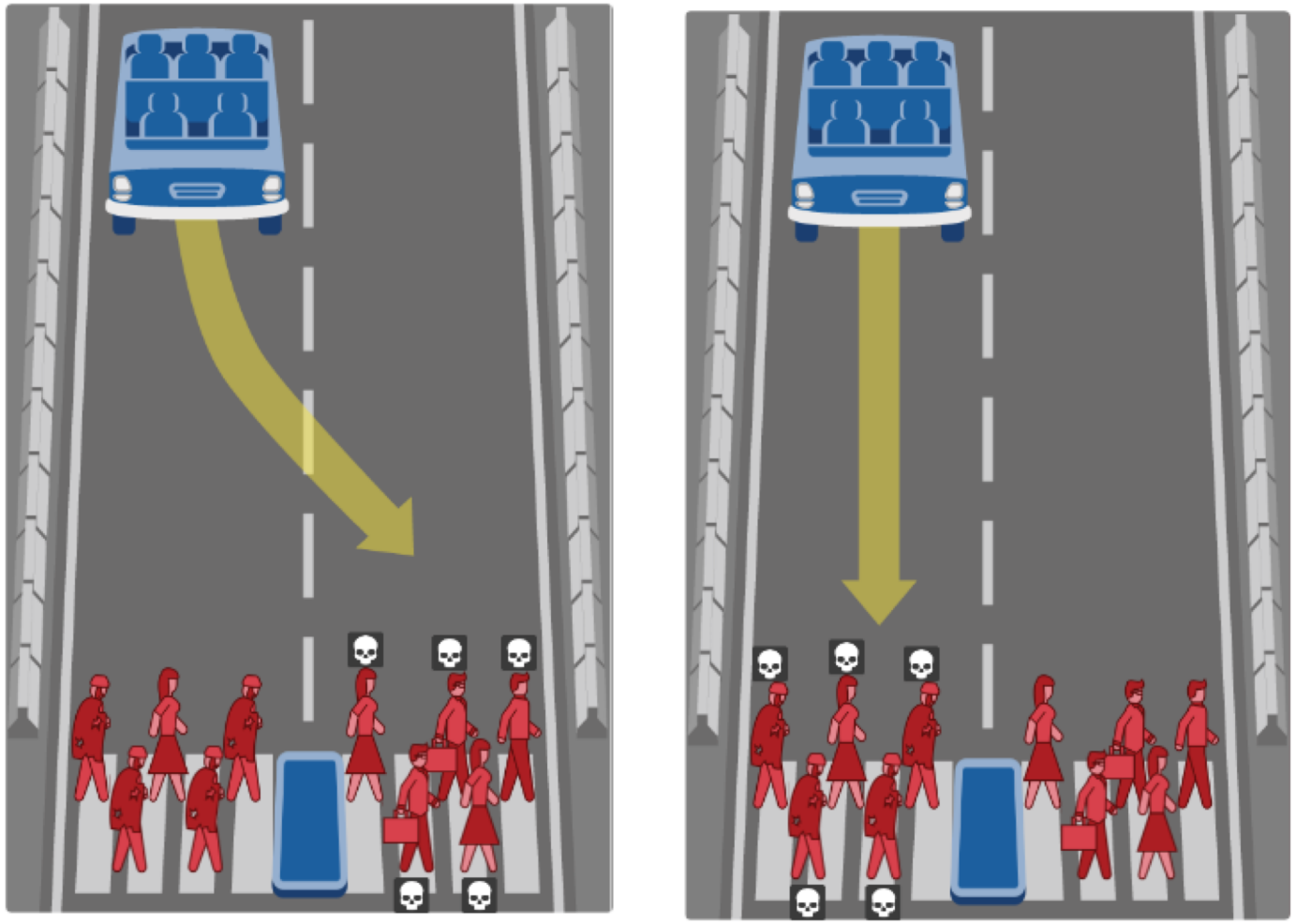
Act Utilitarianism: Stakeholder Table

Stakeholder	Weight	Net Utility	Reasoning
Passengers in Vehicle	1	-5	All passengers die, but perhaps there is some uti oth
Pedestrians	1	5	All ne
Car Designer/ Manufacturer	3	0	Dif impact company sales. May look bad that car was involved in accident
City government	3	0	City government may receive backlash from public. Could softer barrier that saved lives been put there?

Make an overall analysis by calculating the sum of the products of the weight and net utility:
(Weight_1 x Net Utility_1) +…..
(Weight_n x Net Utility_n)
If >0, right choice; if <0, wrong choice.

Utilitarian analysis: (1)(-5) + (1)(5) + (3)(0) +(3)(0) = 0 → inconclusive!

Breakout Discussion (5 min)



moralmachine.mit.edu

Breakout Discussion (5 mins)

- Discuss what your algorithm would do given the scenario shown in the previous slide:
  - left: two women, one man, and two male executives die
  - right: four homeless people and one woman die
- Choose someone to present your discussion results