



NATIONAL RESEARCH
UNIVERSITY

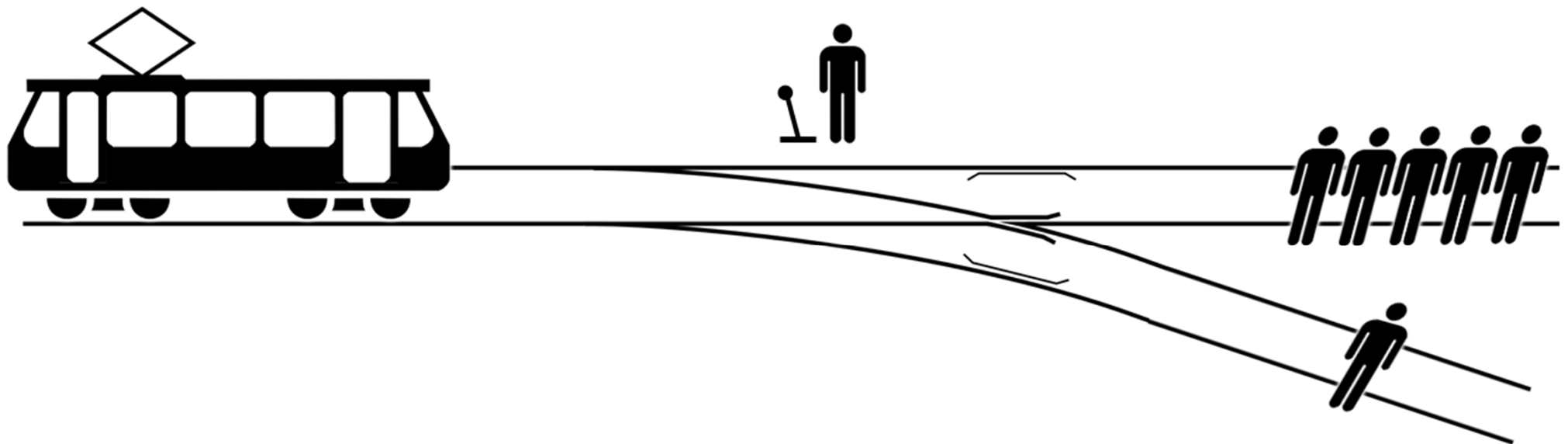
Институт права и развития ВШЭ-Сколково

«ПРОБЛЕМА ВАГОНЕТКИ» И ЭТИКО-ПРАВОВЫЕ ВОПРОСЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Карлюк М.В.

Конференция РСМД
«Международные и социальные последствия
использования технологий искусственного интеллекта»
Москва, 2018

ПРОБЛЕМА ВАГОНЕТКИ: ЭТИКА



coauthor of *WITTGENSTEIN'S POKER*

David Edmonds

**WOULD
YOU
KILL
THE
FAT
MAN?**

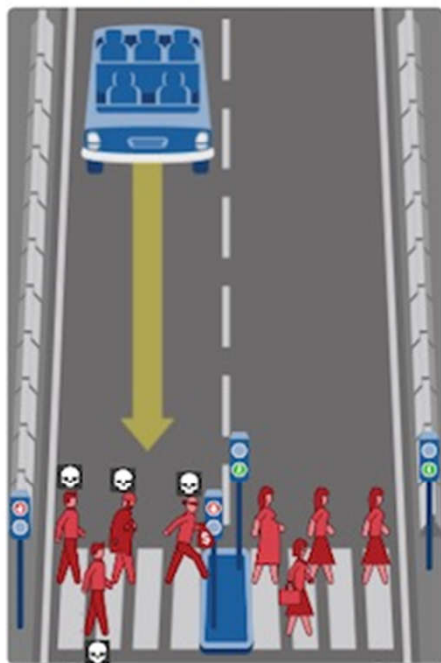


The Trolley Problem
and What Your Answer Tells Us
about Right and Wrong

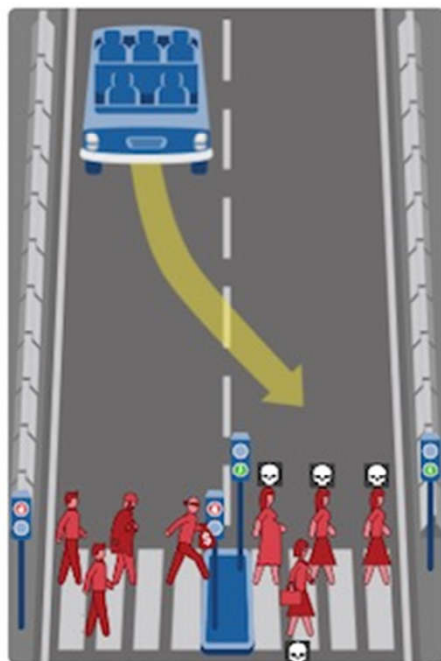
"A tour de force."—Kwame Anthony Appiah, author of *The Honor Code*



What should the self-driving car do?



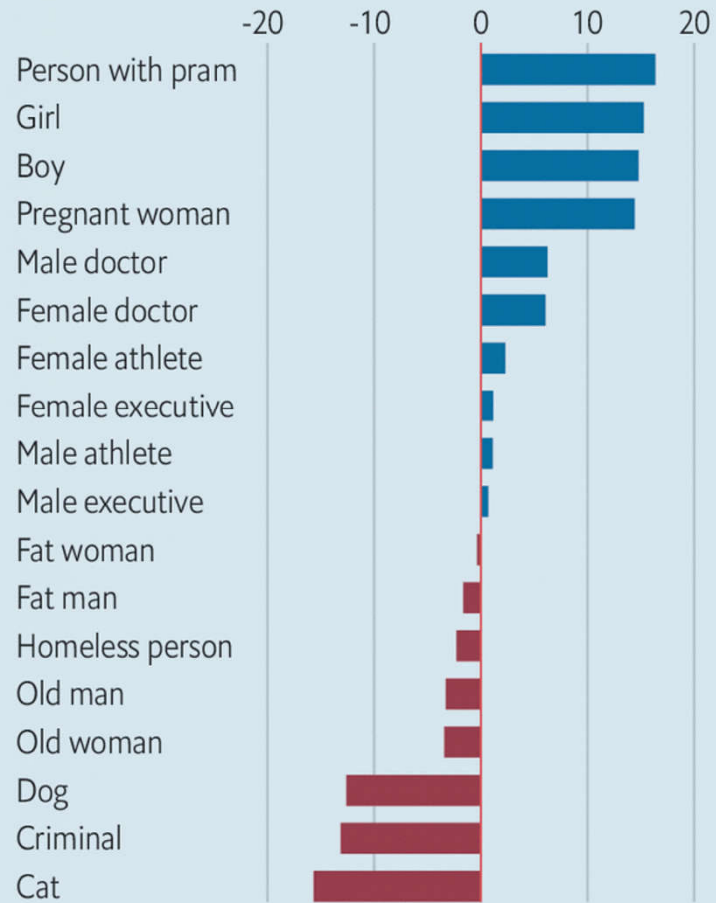
Show Description



Show Description

Hit and miss

Theoretical driverless-car experiment,
respondents' preference of whom to spare
Probability compared with sparing average adult, %



Source: "The Moral Machine experiment"
by E. Awad et al, *Nature*, 2018



ПРОБЛЕМА ВАГОНЕТКИ: ПРАВО

Пример:

- Мерседес
 - машины отдают приоритет пассажирам
- Министерство транспорта ФРГ
 - машина должны отдавать приоритет безопасности человека над причинением вреда имуществу
 - машина никогда не должна делать различия между людьми на основе категорий, таких как возраст или раса
 - если человек отпускает руку от руля управления, производитель машины несёт ответственность в случае аварии



АВТОНОМНЫЕ АГЕНТЫ

некоторые устройства или программы, воспринимающие свою среду и выполняющие действия, которые максимизируют их шансы на успех в достижении определённой цели

Проблемные правовые вопросы:

1. предсказуемость, которая критична для современных правовых подходов
2. способность быть причиной события (действовать самостоятельно), но не нести правовую ответственность



ОТВЕТСТВЕННОСТЬ И ПРЕДУБЕЖДЕНИЕ

- Ответственность разработчиков?
 - мы не знаем, что мы будем знать
 - взаимодействие алгоритмов
- Прозрачность и предубеждение алгоритмов
 - какая информация используется?
 - должны ли быть компетентные органы наделены полномочиями по проверке алгоритмов?
 - кто должен иметь право знать, как алгоритмы принимают решения?



ВАРИАНТЫ РЕГУЛИРОВАНИЯ

1. Автономные агенты = объекты авторских прав
2. Автономные агенты = животные
3. Автономные агенты = источники повышенной опасности
4. Автономные агенты = юридические лица
5. Автономные агенты как автономные агенты?

ГУРУ ГРАНТХ САХИБ, ИНДИЯ



РЕКА УОНГАНУИ, НОВАЯ ЗЕЛАНДИЯ





Te Awa Tupua (Whanganui River Claims Settlement) Act 2017

Public Act	2017 No 7
Date of assent	20 March 2017
Commencement	see section 2

Legal status of Te Awa Tupua

14 Te Awa Tupua declared to be legal person

- (1) Te Awa Tupua is a legal person and has all the rights, powers, duties, and liabilities of a legal person.
- (2) The rights, powers, and duties of Te Awa Tupua must be exercised or performed, and responsibility for its liabilities must be taken, by Te Pou Tupua on behalf of, and in the name of, Te Awa Tupua, in the manner provided for in this Part and in Ruruku Whakatupua—Te Mana o Te Awa Tupua.

15 Legal effect of declaration of Te Awa Tupua status

- (1) This section applies to persons exercising or performing a function, power, or duty under an Act referred to in Schedule 2—
 - (a) if the exercise or performance of that function, power, or duty relates to—
 - (i) the Whanganui River; or
 - (ii) an activity within the Whanganui River catchment that affects the Whanganui River; and
 - (b) if, and to the extent that, the Te Awa Tupua status or Tupua te Kawa relates to that function, duty, or power.
- (2) Persons exercising or performing a function, power, or duty (**decision makers**) under the Acts listed in clause 1 of Schedule 2 must recognise and provide for—
 - (a) the Te Awa Tupua status; and
 - (b) Tupua te Kawa.

European Parliament

2014-2019



TEXTS ADOPTED

P8_TA(2017)0051

Civil Law Rules on Robotics

European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL))

The European Parliament,

- having regard to Article 225 of the Treaty on the Functioning of the European Union,
- having regard to Council Directive 85/374/EEC¹,
- having regard to the study on Ethical Aspects of Cyber-Physical Systems carried out on behalf of the Parliament's Science and Technology Options Assessment (STOA) Panel and managed by the Scientific Foresight Unit (STOA), European Parliamentary Research Service;



ОТВЕТСТВЕННОСТЬ АВТОНОМНЫХ АГЕНТОВ

(Предполагаемые) различия между людьми и автономными агентами (касательно ответственности):

- интенция
- свобода воли

Решение: интенция и свобода волю должны пониматься как то, что мы приписываем друг другу, а не то, что существует в реальности



ВОПРОСЫ

- Какие цели мы преследуем?
- Чьи интересы должны иметь приоритет?
- Насколько эффективно это будет?





СПАСИБО!



NATIONAL RESEARCH
UNIVERSITY

тел: +7 (916) 394 8020

email: mkarliuk@hse.ru