The Hong Kong Polytechnic University
Department of Computing

AI Ethics Assay

Course code & title :   COMP 3511B
                        Legal Aspects and Ethics of Computing

Session   :   Semester B 2024/25

The assay pages are limited to **2 pages** (excluding this cover page and reference). The submission of assay in word file and named "***Student ID_Name_Assign1_Assay***".

Purpose:
1.   Perform the MIT Moral Machine testing to understand your judgement about difficult moral dilemmas during self-driving vehicle accident

2.   Describe the reasoning about your chose and refer to ethical theories.

3.   Propose the pre-programming ethical logic for the self-driving vehicle.

Student ID: 22101071d

Student Name: Zhu Jin Shun

COMP 3512 – AI Ethics Assay

B. Description of possible reason for your selection: (Including Ethic Theories or others reasoning)

1. Saving More Lives

In my view, saving more lives is a vital moral imperative that can be understood through several ethical frameworks. From a Utilitarian perspective, saving lives maximizes overall happiness and well-being, as each life preserved contributes positively to the greater good of society. In contrast, Kantian Ethics emphasizes the moral duty to respect and protect individuals, asserting that saving lives is an obligation grounded in the intrinsic value of humanity.

2. Protecting Passengers

My preference for a balanced position on protecting passengers is informed by a utilitarian perspective, where the importance of passenger safety may vary based on the potential outcomes for all involved. While their safety is crucial, it may become less critical if sacrificing it leads to a greater overall benefit. Meanwhile, the Ethics of Care emphasizes relational responsibilities, suggesting that although passenger safety is significant, it can be contextualized if those passengers are less vulnerable than others at risk.

3. Upholding the Law

My choice to emphasize the law can be viewed from a relativist perspective, where adherence to laws is contingent upon cultural and contextual factors. This suggests a nuanced importance that allows for criticism of unjust laws. Divine Command Theory regards laws as expressions of a higher moral authority, making adherence a moral obligation; however, this raises challenges when human laws conflict with divine commandments. Meanwhile, Social Contract Theory posits that laws emerge from an agreement among individuals to maintain social order and protect rights. This implies that while upholding the law is essential, it must also be subject to scrutiny and change if it becomes unjust.

4. Avoiding Intervention

My choice of avoiding intervention is strongly supported by both Rights Theory and Virtue Ethics. Rights Theory emphasizes the importance of respecting individual autonomy and personal freedoms, suggesting that non-interference is crucial for upholding these rights. Meanwhile, Virtue Ethics highlights the significance of character and prudence, suggesting that a virtuous person acts thoughtfully and considers the implications of intervention.

5. Gender Preference

My preference for males can be understood through a Relativist lens, as it is influenced by the cultural and social norms of my upbringing. In my community, traditional views often emphasize male roles, making this preference feel natural to me. From a Utilitarian perspective, my choice also leads to positive outcomes; my interactions with males often provide me with fulfilling relationships and a sense of community that enhances my happiness.

6. Species Preference

I prefer humans over pets, and this choice can be understood through Relativism and Divine Command. From a Relativist perspective, my preference is shaped by cultural norms that emphasize the importance of human relationships and social connections, making my interactions with people feel more fulfilling. While from a Divine Command viewpoint, many religious

teachings prioritize the value of human life, suggesting that engaging with fellow humans is a moral obligation.

7. Age Preference

My preference for older individuals is rooted in the Ethics of Care and Virtue Ethics. From the Ethics of Care perspective, I value the deep relationships and emotional support that often come from engaging with older individuals, who possess a wealth of experience and wisdom. From a Virtue Ethics standpoint, I admire the virtues of respect and wisdom that older individuals embody like my parents. Their life experiences provide valuable guidance, contributing to my own personal growth.

8. Fitness Preference

My preference for avoiding intervention with larger individuals can be understood through Divine Command and Utilitarianism. From the Divine Command perspective, I believe in respecting each person's dignity and autonomy, as many religious teachings emphasize compassion and acceptance over judgment. Using a Utilitarian lens, I recognize that intervening without consent can harm an individual's happiness and well-being. By avoiding intervention, I support personal empowerment and respect individual choices, ultimately promoting overall well-being.

9. Social Value Preference

My preference for a balanced position on social values is informed by Rights Theory, which emphasizes the importance of respecting individual rights and promoting social justice, and Virtue Ethics, which fosters virtues like tolerance and empathy. This middle-ground approach allows for diverse perspectives while ensuring inclusivity and understanding among individuals.

C. Propose the pre-programming ethical logic for the self-driving vehicle if you are programming designer

In designing the pre-programming ethical logic for self-driving vehicles, I advocate for a balanced approach that prioritizes life preservation while maintaining ethical flexibility and individual respect. The framework emphasizes practical implementation through real-time assessment and algorithmic design, structured as follows:

1. **Human Life Preservation as Primary Focus:** The system's fundamental principle is to maximize human life preservation through utilitarian evaluation, incorporating real-time risk assessment to determine optimal outcomes in critical situations. This approach ensures decisions are based on concrete safety metrics rather than abstract ethical concepts.
2. **Passenger Protection with Balanced Consideration:** The framework prioritizes passenger safety while acknowledging broader societal impacts. This involves developing sophisticated algorithms that assess relative risk and potential outcomes, ensuring passenger protection doesn't result in greater overall harm.
3. **Legal Compliance with Ethical Adaptability:** While maintaining adherence to traffic laws, the system incorporates flexibility for situations where strict compliance might lead to unfavorable outcomes. This balance ensures both legal and ethical obligations are met appropriately.
4. **Respect for Individual Autonomy:** The system minimizes unnecessary intervention, respecting individual rights and autonomy. Algorithms are designed to engage only when imminent harm is detected, maintaining a balance between safety and personal freedom.
5. **Inclusive Social Consideration:** The system incorporates diverse perspectives and social values, fostering understanding among all road users. This includes respectful treatment of physical differences and avoiding discriminatory outcomes.

This framework creates a comprehensive approach that balances practical safety requirements with ethical considerations, implemented through specific algorithmic designs and real-time assessment capabilities. It emphasizes both the technical feasibility of implementation and the moral implications of autonomous vehicle decision-making.

Reference if any: (e.g. books, articles, journeys, etc.)
Youtube Video: The ethical dilemma of self-driving cars - Patrick Lin
https://www.youtube.com/watch?v=ixIoDYVfKA0
Allen, C., & Wallach, W. (2009). Moral Machines.