****

**ENGN6250 /​ COMP6250 |** Professional Practice 1

Week 4 **Tutorial Activity**

**Sample Paragraphs**

**Task: Read the following paragraphs in response to the question, complete the table and bring your answers to the tutorial for peer discussion.**

**Question: Critically evaluate implication of bias or morality when developing AI.**

**Response 1**

The continually developing AI will face moral problems finally, when human cannot completely control AI (Misselhorn, 2018). The author also argues that artificial morality is also required for AI system and AI will be a basic moral agent for itself. However, it is not an easy task. As Misselhorn mentions, morality is made by a serious of principle and attitude and human morality may be not appropriate for artificial system. Concreting abstract ethical rules and implementing them into system are two challenges for the cognitive program may be influenced by every new data (Misselhorn). Moreover, the author illustrates that a hybrid approach AI designed for elder care is created by cooperation of computer science, social science and philosophy. Misselhorn conclude that artificial morality makes machines better and may also benefit human civilizing moral capacities (Misselhorn, 2018).

**Bibliography**

Misselhorn, C. (2018). Artificial morality. Concepts, issues and challenges. *Springer Science+Business Media, 55*, 161-169.

**Response 2**

Autonomous vehicles bring convenience and benefits for people, but there are some moral issues that the self-driving technology may run into.

The basic capacity of an automated system is to make right judgement about all the signals and messages it may receive on the road, which means sometimes dealing with some fuzzy signals. For example, a vague data or signal should be distinguished correctly that it fundamentally is a sensor noise, sunlight reflecting from the road or a little child. Additionally, driving is an interactive activity between people and people or people and objects, so it involves a series of signals and messages from simple gestures to some unspoken rules. However, based on some flawed or omitted data algorithms of autonomous vehicles may make incorrect judgements, because the computer vision algorithms which may have innate bias on analyzing the circumstances along the road are the main resource to support self-driving technology (Howard & Borenstein, 2019, p.3). According to Amitai and Oren (2017), scholars suggested that self-driving vehicles should have ability on estimating whether the decision is wrong or right as sometimes these vehicles may cause harm to both passengers and pedestrian. More than that, Misselhorn (2018) points out “Another important field of artificial morality is autonomous driving which poses particular difficulties because autonomous vehicles do not just face moral decisions but moral

dilemmas.” The research by Foot (2002) illustrates a well-known example trolley problem, which is a test of moral issue about whether it is morally permissible to save several people’s lives even one person’s life is required to be sacrificed.

In conclusion, some moral issues need to be considered before autonomous vehicles are widely manufactured and used in society.

**Reference lists**

Etzioni, A., & Etzioni, O. (2017). Incorporating ethics into artificial intelligence. The Journal of Ethics, 21(4), 403-418. doi:10.1007/s10892-017-9252-2

Foot, P. (2002). The problem of abortion and the doctrine of the double effect. (pp. 19-33) Oxford University Press. doi:10.1093/0199252866.003.0002

Howard, A., & Borenstein, J. (2019). Trust and bias in robots. American Scientist, 107(2), 86-89. doi:10.1511/2019.107.2.86

Misselhorn, C. (2018). Artificial morality. concepts, issues and challenges. Society, 55(2), 161-169. doi:10.1007/s12115-018-0229-y

**Response 3**

With the development of Artificial Intelligence (AI), how AI should make moral decisions has been a big problem for scientists designing AI systems. Etzioni and Etzioni (2017) analysed ethical problems when AI makes decisions using ‘the Trolley narratives.’ Two approaches called ’top-down’ and ‘bottom-up’ were proposed to design an ethics AI system. This also lead to questions about whether human should trust robot decisions in similar situations like driving a car. Howard and Borenstein (2019) suggested a strategy to reduce the influence of problems like over-trusting AI when humans may ignore the risk of giving robots authority to make decisions. Furthermore, it is a problem for AI to deal with problems such as whether a vacuuming robot should kill a spider or a small bug on its cleaning way (Misselhorn, 2018). These moral problems are also a challenge for human beings. Without finding ways to design AI to tackle such dilemmas we may find that AI systems may not be trusted to be fair and neutral in directing human attention and triggering certain actions (Bossmann, 2016). When designed right, AI could evolve into an opportunity to encourage society towards more beneficial behaviour. However, in the wrong hands and influenced by bias it could prove detrimental. It is necessary for scientists to further explore AI systems as a moral agent and tackle resultant implications when developing AI.

**References**

Bossmann, J. (2016). Top 9 ethical issues in artificial intelligence. World Economic Forum. Retrieved from https://www.weforum.org/agenda/2016/10/top-10-ethical-issues-in-artificial-intelligence/

Etzioni, A., & Etzioni, O. (2017). Incorporating Ethics into Artificial Intelligence. *The Journal of Ethics*, 21(4), 403–418. doi: 10.1007/s10892-017-9252-2.

Howard, A., & Borenstein, J. (2019). Trust and Bias in Robots. *American Scientist,*107(2), 86-89.

Misselhorn, C. (2018). Artificial Morality. Concepts, Issues and Challenges. *Society*, 161–169.

*Response were Adapted from student work PP1 S2 2019*.

Based on the following responses, complete the following table.

|  |  |  |  |
| --- | --- | --- | --- |
| **Activity** | **Response 1** | **Response 2** | **Response 3** |
| Appropriate answer to the question |  |  |  |
| Topic sentence |  |  |  |
| Conclusion |  |  |  |
| Coherency/Logic in arguments |  |  |  |
| Critical thinking |  |  |  |
| Quote, if present, done appropriately |  |  |  |
| Missing intext citations |  |  |  |
| In text citations mechanics |  |  |  |
| Reference list- compatible with IEEE/APA/Harvard |  |  |  |
| Consistency in reference list |  |  |  |
| Comments on sentence structuring |  |  |  |
| Likely grade |  |  |  |