**Philosophy Paper**

**Word count: 1419**

The famous series of movies Avengers is known for its imagination, among which the movie Age of Ultron is about AI technology. The AI Ultron is first designed aiming in protecting the earth and maintaining peace. But since its developers Tony Stark and Bruce Banner did not build well-designed ethic restriction in Ultron’s base logic, it soon fell to the opposite--considering human as the main factor to the Earth and peace. This turned out to be a disaster at last. This essay will focus on the ethic dilemma shown by Ultron(the AI technique this movie contains), some conception of real world solution, as well as some benefits of applying AI technique.

One ethical dilemma it pose is that, if AI not be under proper ethic restriction, it may generate different understanding to human instructions and cause problems. This is reasonable because AI learns from data sets(to some extent also human do) which can cause bias, like we discuss in the lecture(Lecture, 22/9) about AI. In the movie, Ultron learnt from his previous attempt of maintaining world’s peace. Every time he witnessed what human did to cause Chaos, he strengthened his recognition on “human is the biggest problem to peace”. It’s easy to see that it’s difficult for Ultron to see anything else but human-caused Chaos since he is designed for solving them and to some extent human is the biggest problem maker of all species. And reaching Ultron’s conclusion is probably because of incomplete data set, which can be solved by proper ethic restriction. But on the other hand, it is really hard to illustrate ethic standards to AI since ethic standards of human is still not a clear complete framework. Some of human’s main values even contradict to each other. When we discussed equality in the lecture session previously(Lecture, 10/11), we stated that it is important to make weighting between different core values since it is almost impossible to reach them together. Just like the movie, there was no clear definition for “peace”: for Tony Stark and Bruce Banner, peace may means “most moral people living under protection of law”, which is a stage need to be maintain; But for Ultron, all humans are considered to have potential to pollute peace, and peace is a “constant” that can’t be change. In other words, humans are all potential factor able to change this constant. The only way to finish his task is to eliminate all humans.

Thoroughly, both AI’s understanding and human’s understanding of peace may be true to some extent. But more importantly, taking AI as our tool, it will be better if we make AI “our side”.

In the movie, there’s no such implementation by laws or ethics, Ultron’s obedience was made by immortals. In real life, there also is no such restrictions since AI is not that advanced at current time. But there is some ideas in public that AI should follow rules like ”the three law of robots”, which was introduced by famous sci-fi writer Adolf Kosimov. For this solution, there are some challenges too. For example, It will short circuit AI when solving trolley problem based on the principles. The key is that no matter what ethic principles is given in specific definition, it will meet its limit in infinite decision making. In other words, ethics contains things over definitions which require further implementation.

In my opinion, a possible solution is to let AI learn ethics like human does. In the design, AI should first have some common sense of core values. Then it update its recognition of these values by machine learning. The most important part is human should frequently communicate with AI about these recognition such that they can reach agreement. This is like doing therapy for AI to correct it from misunderstanding of core values.

Based on this method, we won’t need laws to govern AI since it always take a similar side with human instructors. But we do need to form some norms around AI.

One of the most important things is that there should be IT and ethical professionals checking AI’s recognition time by time. Like we discussed previously, in the healthcare domain, AI systems are used for assisting in diagnosis and treatment decisions. To ensure the accuracy and ethicality of AI in this field, it is crucial to have dedicated healthcare and ethical professionals periodically reviewing the AI system's recognition results. These experts can assess the performance of AI across different diseases and patient populations, providing feedback and suggestions for improvement to ensure the correctness and reliability of the AI system. Besides, the importance of transparency and explainability should not be overestimated. AI should allows users and stakeholders to understand how the system makes decisions to avoid hidden incorrect recognition. For example, in the financial sector, AI systems are widely employed for risk assessment and credit scoring. To prevent hidden incorrect recognition, transparency and explainability are of utmost importance. Users and stakeholders should be able to comprehend how the AI system arrives at decisions, including the data used, algorithms employed, and factors considered. This enables them to scrutinize and validate the AI system's decisions, ensuring accuracy and fairness. If the decision-making process of AI systems is excessively opaque, making it difficult for users to understand the underlying logic, it becomes challenging to identify and rectify potential errors or biases. Also it is highly suggested to build the norms under global sight. Consider the situation of movie Ultron, the AI being powerful and able to create huge disaster largely credits to full access to the whole world. If there is no global norm and AI is managed differently in different district, some potential risk will be exist in data transmission that AI in a district may not satisfy ethic requirements of another district.

If above measures are correctly implemented, there’s still a final question of AI: “Is it worthy?” We spent many time, money and manpower to simulate human thinking, massively producing such simulation and “legalize” it. The answer is a certain “yes”. Applying AI technique has several benefits.

The first benefit is that AI provides considerably efficiency and scalability compares to human thinker. It offers the potential to process vast amounts of data, perform complex calculations, and make decisions at a speed and scale that surpass human capabilities. This efficiency can lead to significant advancements in various fields such as medicine, finance, and transportation. For example, AI-powered medical imaging systems can analyze thousands of medical images in a short time, assisting doctors in diagnosing diseases more accurately and efficiently. Beside, AI is also a good choice in repetitive and dangerous tasks.By automating these tasks, AI frees up human resources to focus on more creative and complex problem-solving. Also, AI can be deployed in environments like deep-sea exploration or space missions, where human presence may be risky or impractical. In addition, AI has the potential to augment human intelligence and decision-making rather than replacing it entirely. By leveraging AI as a tool, humans can access and process vast amounts of information and make more informed decisions. For instance, AI-powered recommendation systems can help individuals discover new books, movies, or products based on their preferences, expanding their knowledge and enhancing their experiences. Moreover, the combination of human intelligence and AI capabilities can lead to novel solutions and discoveries. By collaborating with AI, humans can leverage its computational power and data analysis capabilities to tackle complex problems and drive innovation across various domains. The last but not least, AI may provide a new aspect for human seeing themselves. As in the movie Ultron, AI stands from a different angel with humans. Although the viewpoint is common, but what matters is that it illustrate exciting potential that AI may help break human recognition limit. It may be able to rapidly analyze and learn from extensive data sets, potentially uncovering patterns and insights that would be difficult for humans to detect. Comparing to these unique benefits, AI’s high marginal cost: developing and maintaining is acceptable.

In summary, AI, although showing potential ethic risks, when properly guided by ethics and human interaction, can be a powerful tool that accelerates progress in various fields. By implementing effective norms and measures, we can harness its capabilities while ensuring its alignment with human values. AI has the potential to revolutionize industries, break cognitive barriers, and offer valuable insights that humans alone may struggle to uncover. With careful consideration and responsible implementation, AI can undoubtedly contribute to more labour and more advanced recognition for humanity.

**Pre-writing:**

**Outlining:**

Who(will consider): policy maker; user(individual and company)

What: should AI be self-determined or to what extent should we restrict it;at the end we go back to moral of mankind.

When(will it be possible): probably not far away from now

Where: everywhere with AI technology(in the movie is everywhere)

Why:because we want AI do intelligent works for mankind which requires many decision making

**Clustering:**

Main idea: AI should be trained for ethics and restrict use from a policy standpoint. There may not be laws to govern it since the inner truth of AI(acting as human) makes it meaningless to restrict how it implement. But it is possible to form some norms around the use of it. For example, AI should pass a recognition test frequently. The technology it self should exist since it provide huge convenience if used correctly

Brief Introduction to the movie

Analysis of ethical dilemma it pose

Exact policies: no

Some current solution and disadvantage

How I design to treat AI and advantage

**Reflection Essay**

**Words count:508**

As previously discussed in the lecture(Lecture, 10/11), diversity and inclusion offer diverse sight for destructing problems. So my design for company regulation to strengthen diversity will be as follows.

First, diversity and inclusion should be a core value in our company introduction. I hope this can ensure that every one of our employees knows our commitment to diversity and inclusion, and sees them as the core factor of implementing our company goal.

Besides, applying diversity and inclusion strategies in hiring should be carefully considered. On one hand, it is important to take positive measures to attract and hire people from different backgrounds, culture and experience. This can be implemented through building cooperation relationship with diverse organization and societies. But on the other hand, it will be inappropriate if we overvalue diversity in hiring standards. In all, equal standards that eliminate bias to different groups of people will be what we need.

Inside the company, we will have several arrangement to ensure diversity and inclusion:

First we will encourage our employees to communicate with people from different backgrounds, culture and experience. We will have a permanent event like opinion sharing day on every Friday afternoon. It will be like a party but all participants are encouraged to share their unique points of view. This event aims at promoting idea sharing, and we hope our employees can see different ways others consider in and value diversity.

Also, I will include lectures on diversity in our employee education plan. All employees are encouraged to take these lecture to get fully understanding to diversity and inclusion. We may also invites celebrities from different backgrounds to visit our company. Through these measures we hope our employee can get a deeper sight into diversity and realize how deeply it rooted in our life.

For company policies, we highly value and integrate diversity and inclusion into our decision making. When consider employee managements like promoting, we are more likely to offer chance for those who show respect to diversity and inclusion. Also, we will be equally serious to any demand from any one of our employee. We will take positive attitude to consider our employee’s demand.

When talking to the leader team, it will be consistent about our diversity and inclusion policies, that is we will work hard to build a diverse leader team. In such way we want to ensure voice from different backgrounds, cultures and experience can be involved and considered in decision making process by equal chance, which will help in building a more diverse and equal working space for all members. It will also provide more chance and developing space for every employees.

Finally, a timely and frequent evaluation process should be maintained. We will evaluate our company’s diversity and inclusion, which will be quantified, and planning further improvement based on the result. To make it specific, it may contain process like collecting employees feedback, research in finding employee composition and corresponding data analysis. By doing so, we hope our company can always take diversity and inclusion into good practice.

**References**

Zhicong, L. (2023, November 10). JC2066: IT Professionals: Ethical, Legal and Social Issues L07: Inequality and Justice in Computing [pptx]. <https://canvas.cityu.edu.hk/courses/55183>

Zhicong, L. (2023, September 22). JC2066: IT Professionals: Ethical, Legal and Social Issues L03: Artificial Intelligence [pptx]. <https://canvas.cityu.edu.hk/courses/55183>