"The Moral Machine Experiment"

Summarize the key questions the paper aims to address:

The key question discussed in the paper is "What are the defining factor in the development of global principles for machine ethics?" The specific sub-questions include,

- How global moral preferences can contribute to the development of socially accepted principles for machine ethics?
- The role of individual preferences on the creation of universal socially accepted principles for machine ethics, like in pedestrians and moving objects.
- Effect of cross-cultural ethical variations in the development of socially accepted principles for machine ethics
- How machine will make moral decisions in today's world and should guide machine behaviour

Summarize the method(s) used in the paper:

The methods used in the paper,

- Hierarchical clustering analysis involves finding groups within a sample that align together. In this paper, the clusters were scored based on Euclidean distances.
- Regression analysis where each of the nine attributes was regressed on the respondent's final decision.
- Correlation analysis was used to find the relations between moral preferences extracted from the Moral Machine and the cultural distance from the US. Specifically, the ordinary least squares method was used to show the relationships between key ethical preferences and various economic, political, and social measures.
- Descriptive statistics were used to estimate the averages and standard deviations of the results given by the respondents.
- The average marginal causal effect (AMCE) method was used in combination with conjoint analysis.

Summarise the key results from the paper:

Under global preferences – strong preferences were shown on sparing humans compared to animals, sparing young lives and more lives compared to the old and fewer individuals.
The study results showed that globally accepted moral ethics should solve the following dilemma as follow:

- Choose between more or fewer people (answer sacrifice few).
- Sacrifice between the young and the old (answer sacrifice the old).
- Sacrifice an animal or human being (answer spare human being).

Under individual variations – Both the major gender (male or female) concur that females should be spared when an autonomous vehicle is faced with a dilemma of choosing between male and female. However, the preference for sparing females is stronger if the respondent is female.

Under cultural clusters - The preference of sparing the younger generation compared to the old is less noticeable for countries in the east than other regions. The intensity of sparing the old becomes more pronounced for countries in the south. Similarly, east countries reported sparing people with high status than low-status individuals. Further, countries in the southern and eastern clusters reported a high preference for sparing pets to human. However, all the clusters concur that the lawful should be spared than the unlawful.

On country-level predictors – countries with higher inequality corresponded to the varied treatment of two classes of people. That is people from fewer equality countries treat the rich differently in the Moral Machine.

Discuss if the methods and results presented are robust in your view:

The method of data collection is robust since a sample size of more than 39 million globally is sufficient to generalize. There are variations in dilemma solving based on the region, demographic factors, individual preferences, and economic status. This kind of differences is best tested using hierarchical clustering. With this method, the researcher can identify patterns in the response while considering the differences in other factors. The paper used basic statistical techniques such as ordinary least squares regression (OLS) which operates under several assumptions. In the entire paper, the assumptions were not tested yet some of them may have been violated. In this regard, the results are not robust in my view.

Identify and explain 2 potential positive effects on society that could result from the work presented in the paper:

Initially the society will benefit from the improvement of services since the machine is efficient. Now that some of the profound moral ethics have been brought to the limelight to this paper, the discussion on possible and generally accepted solutions will go on. The discussions will result in the development of machines that coexists with human better. This paper is like an eye-opener to the development of machine moral ethics.

Also, society will benefit from the incorporation of moral ethics in their machines in that they will reduce the chances of erratic occurrences presented by humans when using machines. For instance, in a country where younger generations are more valued than the older generation, accidents involving younger generations will be avoided. Thus, societal objectives will be achieved in the long run.

Identify and explain 2 potential negative effects on society that could result from the work presented in the paper:

First, the paper seems to be in support of the mass adoption of machines- Which will result in loss of employment to the general populations. Loss of employment will mean that society will face negative effects of low employment especially in developing countries that heavily depend on labour compared to capital. Also, if any issue will happen technically, there might be high chances of failure.

Second, the successful incorporation of machine ethics will increase trust in the machines and result in overreliance. Thus, some of the technical skills like driving cars will become obsolete.

Propose a new experiment/analysis that could be carried out to improve the work presented in the paper, giving details of how you would change or replace the data collection and/or analysis techniques used, and explain why it would be superior:

In the new experiment, the same data collection process will be used, however, changes will be made to the method of analysis of the data.

- The first key factors will be identified using exploratory factor analysis followed by principal factor analysis.
- Second, logistic regression will be used instead of multiple linear regression. The use of logistic regression will show the probability of choosing between two outcomes. However, in cases where the outcomes are more than two, multinomial regression will be used.
- Third, the paper failed to test the assumptions underlying the methods use; in this new experiment key assumptions will be tested.
- The hierarchical clustering still will be used and compared with the results obtained from logistic regression.
- The results will be robust since appropriate transformations will be applied to the data to make it conform to the assumptions underlying the methods used.
- Further, the test will be used to check for the significance of the noted differences. For instance, ANOVA might be used to check the general fit of the regression model. Finally, I would suggest performing a systematic sampling of the original sample so that to improve the representativeness of the sample.

Sharing my real time experience in Google Waymo Project, when I was working at Google, the project is related to machine learning (Self-Driving Car (SDC)), So we have seen the real time data how machine can see and understands it. So, we have generated categorized labels manually to understand SDC much better in difficult scenarios. for example, identifying human and moving objects. So, we analysed and given instructions to the SDC by using censor for real time data as required.

So, proposing the idea is like when will do with the real time data, that would be more chances to understand the machine in moral cases.

Briefly identify and estimate what human, computational and other resources would be needed for your proposed new experiment/analysis:

The data collection will be online so the experiment should be hosted in interlinked servers in every continent. This will ensure that the servers do not surfer overload which may result in incomplete data. Data analysis experts will be needed to help with the validation of the proposed methods. Additionally, there should be intense marketing of the Machine Moral questionnaire to ensure that the sample size increases and becomes more representative. For instance, Africa was underrepresented, while North America and Europe were over-represented.