Human-Machine Cooperation and Conflict: the relationship between safety consciousness and behavior intention of self-driving for seniors from a non-linear relationship perspective

Yu-Kai Huang, Associate Professor
Bachelor Degree Program in Ocean Tourism Management, National Taiwan Ocean University

1. Summary

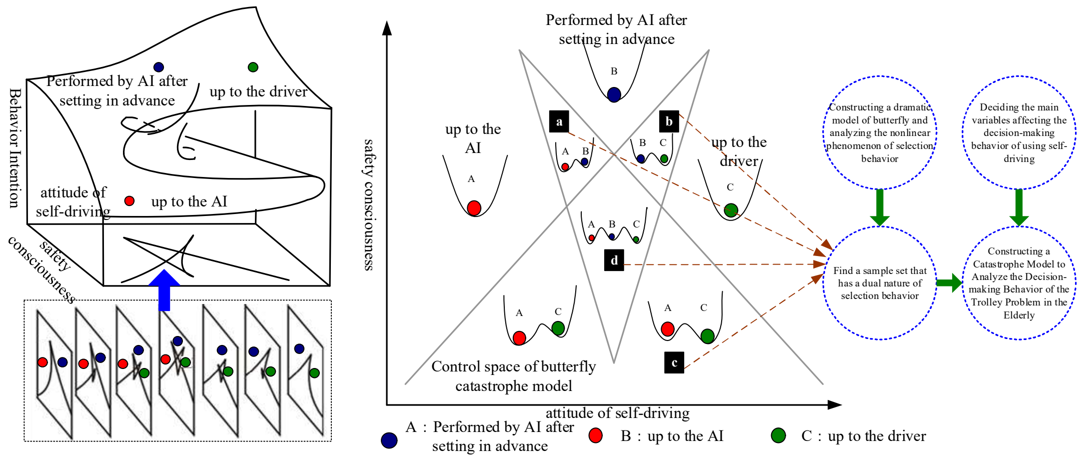
Artificial intelligence will inevitably dramatically change the way people live. In this wave of artificial intelligence changes, “humans can't keep up with the speed of AI technology changes, and they can't adjust and adapt in time” may be more important issues than the development of AI technology. The development of road transport from today's assisted driving to future autonomous driving will be an inevitable trend. At present, we have entered the Level 3 stage of self-driving by assisted driving; therefore, it is an important research topic for exploring safety consciousness and the behavior intention of self-driving for seniors. First, this study analyzes the non-linear relationship between the safety consciousness and the behavior intention of self-driving based on the butterfly catastrophe model on the elderly. Secondly, we construct the “Trolley Problem” situation using self-driving, analyze human-machine cooperation and conflict time. The results found that when autopilot is activated and encounters a recognizable chasing situation, most drivers will choose to hand over the driving control to AI; as for those who choose to use AI to fulfill prior driving decisions, the decision mostly chooses to reduce casualties, but the result of reducing casualties actually leads to harm to passersby who obey the traffic rules. Our findings indicate that ethical issues are inevitably involved in human-robot cooperation and conflict, and also make us aware of the implications of re-examining traffic regulations and human safety.

Keywords：Self-Driving, Artificial Intelligence, Moral Dilemma, Human-Machine Cooperation

2. Aim of Research

In recent years, road transportation has entered conditional driving automation (Level 3) from partial driving automation (Level 2). In the future, the relationship between drivers and Level 3 or Level 4 self-driving cars will be “human-machine cooperation〞. Humans and machines “can” and “must” be close to work together to complete the driving task. The application of artificial intelligence will inevitably affect the way people live and work, and self-driving is an important scene of AI application in the field of transportation. In the process of the progress of artificial intelligence, the topic of how people and machines work together will be an interesting and important research topic. This project will use the elderly as the analysis object to construct the behavioral model by using the butterfly dramatic model to analyze the behavioral intentions of the elderly using automatic driving, and the driving decision behavior in different simulated situations. Figure 1 is the research structure of this study, and the work expected to be completed in this project has the following two points:

1. Using the butterfly dramatic model to analyze the behavior patterns of elderly people choosing automatic driving.
2. According to different simulation scenarios (such as tram dilemma), discuss the decision-making behavior of senior driving for these different situations, and propose management issues related to traffic management based on the analysis results.



Picture 1 Research architecture

3. Method of Research & Progression

Cusp catastrophe theory was developed and popularized in the early 1970's. The theory seeks to characterize phenomena where small variations in independent variables may cause large variations in a dependent variable. Thomas believes that as long as the system is observed to have any feature of the dramatic change model, the system to be studied can be described by the theory of dramatic change. Based on the literature on the application of the dramatic change model, we sort out the four steps of the analysis process of the dramatic change model. as follows:

(1) Observe whether the system to be studied has the characteristics of a drastic model. In the eight features of the dramatic change model compiled by Gilmore, if the control variables of the system fall into the set of divergent points, five of them can be observed. The equation can observe three other features. When we find that the system has any dramatic change characteristics, we can assume that the system is suitable for using the drastic model to perform qualitative or quantitative analysis of its nonlinear behavior.

(2) Select a suitable set of control variables and state variables and determine the dramatic model based on the number of control variables and state variables.

(3) Decide which of the control variables are regular factors? What are the splitting factors?

(4) According to the characteristics of the collected data, select the appropriate parameter estimation method for the drastic model: the time series data is suitable for evaluation by the Guastello method, and other data types can be evaluated by Cobb or GEMCAT.

In mathematical terminology, the three dimensional phase space of the cusp model can be described by the following potential equation:

$$grad\ f(x,c) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right) \tag{1}$$

The dynamic system of the above formula can be rewritten as follows:

$$\frac{dX}{dt} = -grad\ f(x,c) = -\frac{\partial f(x,c)}{\partial x_i},\ i = 1,2,\ldots,n \tag{2}$$

If the system state variable X can approach and maintain the same position under long-term changes, then X is said to be the equilibrium state of S, expressed as e_i:

$$e_i : \frac{dX}{dt} = -grad\ f = -\frac{\partial f}{\partial x_i} = 0, i = 1,2,\ldots,n \tag{3}$$

Figure 2 is a diagram of the results of the analysis of the drastic model using the example of the R software. The analysis of the catastrophe model allows us to determine which variables in which situations will produce discontinuous changes in the behavior of the system, and the dramatic model can also Help us to group the samples and explore the intrinsic reasons for the nonlinear characteristics of the system according to the characteristics of different groups.

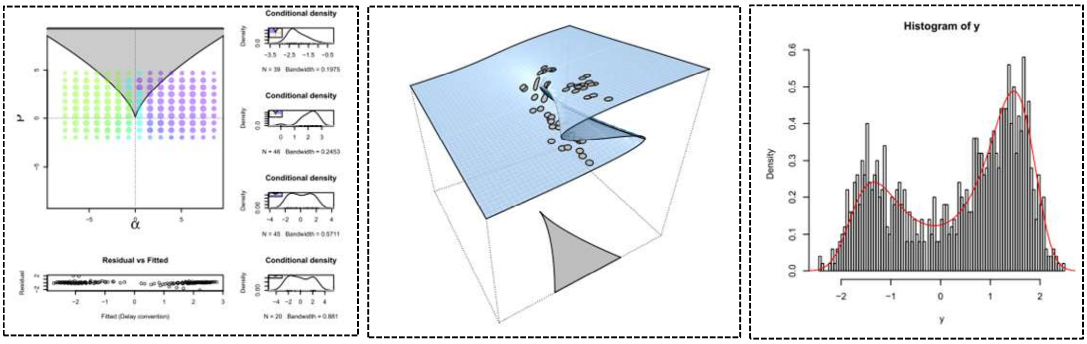


Figure 2 Analysis of the dramatic change model using R

4. Results of Research

The development of self-driving cars is the most eye-catching innovative application of artificial intelligence in the transportation system. In the future, there will be Level 5 self-driving cars based on artificial intelligence in road transportation. Because humans cannot drive Level 5 self-driving cars, the purpose is to explore what kind of ethical decisions we want self-driving cars to show when Level 5 self-driving cars encounter moral dilemmas. Since many documents point out that Level 5 self-driving cars are expected to be introduced into the market between 2025 and 2045, and will reach a 50% market share between 2035 and 2050 (Litman, 2017; Nieuwenhuijsen et al., 2018), this article takes the current age of 46 Ethnic groups over the age of the year are the object of analysis, and the two studies constructed in this article are analyzed. The following explains the main findings of this article:

Study 1 analysis when a self-driving car encounters a choice between sacrificing one self-driving car owner or two pedestrians on the road, what are our expectations for self-driving cars? The results of the analysis found that when the owner and two passers-by are both 30-year-old male office workers, or the two passers-by are friends of the interviewee’s acquaintance, 56.3% of the interviewees said that the goal should be to reduce the greatest human casualties. Therefore, one vehicle owner should be sacrificed to save two pedestrians on the road. However, when a pedestrian violates traffic rules, or when there are relatives in the car, 60.7% (passers-by in violation) and 40.9% (with relatives in the car) are respectively affected. The interviewer believes that self-driving cars should choose to protect the people in the car, even if this result will increase human casualties.

Demographic characteristics such as gender and age have some statistically significant views on how self-driving cars should be determined. Among them, female samples are more likely to be inclined to "passers-by violations" and "family members in the car" than male samples. Decision-making is handled by a random method, and the 56-65-year-old sample compared with the 46-55-year-old ethnic group when faced with "passers-by violations" and "family in the car", more proportions believe that self-driving cars should choose to save passersby (Two people) instead of the owner (one person). When choosing random, it means that the process of "random" does not consider whether the decision is made by "people" or "AI", and there is no so-called objective function for "random", so it can avoid the problem of determining the attribution of responsibility afterward. The results of the analysis show that female samples tend to be more likely to make decisions in a random manner when faced with "passers-by violations" and "family members in the car" than male samples. In addition, according to the analysis framework of the butterfly catastrophe model, the emergence of "random" decision-making means that it is in the early stage of the development of Level 5 self-driving cars, and people generally distrust the decision-making ability of AI.

Study 2 is to explore the moral dilemma when a self-driving car is running empty and is going to pick up its owner, because of the brake failure, it is necessary to choose between two pedestrians to save one person. The results of the analysis show that when the self-driving car faces the moral dilemma of choosing one of the two, Age is the most prioritized variable, and other variables are, in order of importance, the importance of the passer-by in the social network, nationality, whether it violates the rules, and finally gender. The results of the joint analysis show that regardless of gender or age if the passer-by has the following characteristics, he will be regarded as the object most in need of rescue. These characteristics are “0-14 years old”, “high social importance”, “Nationality”, “No violation” and "Female”.

5. Future Areas to Take Note of, and Going Forward

As AI technology enters our daily lives at a faster and faster rate, people increasingly need AI systems to work with humans, which requires AI systems to exhibit behaviors that humans can explain (Kambhampati, 2019). On the other hand, if more and more people choose self-driving cars with Level 3 functions and start automatic driving in the context of allowing automatic driving, the traffic pattern at this time may be due to "human driving" and conditional "automatic driving". "Driving" also exists in a certain section of road, which has an impact on the micro car following behavior and macroscopic traffic patterns. This type of impact may make road traffic more efficient, but it may also cause accidents (Chen Liwen, 2018). The literature points out that when AI is applied to the transportation system, there are many issues that we need to solve. The ethical issues involved in the application of self-driving cars have come. When we do not make decisions, we will be forced to make decisions for us by machines. The question is whether AI’s decision-making is in line with human ethics (such as saving a child) Still an old man, do you want to save two or one people, even if the two are violating traffic rules), if we don’t trust AI's decision-making, can people make better decisions than AI? These ethical dilemmas without standard answers will be unavoidable problems that self-driving cars that act as smart transportation systems in the future will not be avoided. In addition, are the ethics and morals of different countries or races consistent? The standards and definitions of ethics and morals will influence decision-making. Technically, it is the training of artificial intelligence, which appears to be the behavior of self-driving driving. Does the logic establishment and training differ from car owners? This is a matter of the attribution of responsibility for the accident, and it is worth discussing. Therefore, it is recommended that follow-up researchers can conduct a more in-depth analysis of the ethical issues of self-driving cars.

6. Means of Official Announcement of Research Results

The research content of this project has been published in the following journals:

(1) Yu-Kai Huang (2020), “A Preliminary Study on the Moral Dilemma Decision Behavior of Level 4 Self-driving: an Applied Trolley Problem,” *Journal of the Chinese Institute of Transportation*, 32(4), pp. 1-36. (TSSCI)

(2) Yu-Kai Huang (2020), “Human-Machine Cooperation and Conflict: the relationship between safety consciousness and behavior intention of self-driving for seniors from a non-linear relationship,” *Taiwan Insurance Review*, 37(1), pp. 1-33.