AUTOMATED SCENARIO GENERATION FOR SAFETY COVERAGE AND TESTING

(Technical Paper)

AN ANALYSIS OF ETHICAL AND MORAL STANDARDS IN RELATION TO AUTONOMOUS VEHICLES

(STS Paper)

A **Thesis Prospectus Submitted to the**

Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements of the Degree
Bachelor of Science, School of Engineering

Jaspreet Ranjit
Spring, 2020

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Signature Jaspreet Ranjit Date 08/06/20
Jaspreet Ranjit

TECHNICAL ADVISOR
Madhur Behl, Department of Computer Science

STS ADVISOR
Toluwalogo B. Odumosu, Department of Engineering and Society

Introduction

Autonomous driving systems have machine learning components, such as deep neural networks, for which formal properties are difficult to characterize. It is difficult to characterize all of the behaviors of these components under all circumstances. Due to the rarity of failure events, real-world test driving alone cannot provide high confidence in the safety of automated driving systems with respect to injuries and fatalities (Wakabayashi, 2018). This leads to a challenging issue today for automated vehicle manufacturers and suppliers who are determined to incorporate machine learning for automated driving.

Technical Project

The aim of this research thrust is to propose a new innovative certification scheme allowing to demonstrate the level of safety and reliability which allows for safe market introduction of automated/autonomous vehicles. Our goal is to answer the following questions:

Q1: How can we fairly compare two different AV software stacks on a given safety metric?

Q2: How can we leverage simulation to find edge cases and failures for a given AV system?

The kind of closed-loop verification likely to be required for AV component testing is beyond the reach of traditional test methodologies and discrete verification.

Current safety standards for AVs

The meaning of safety in regard to AVs is surprisingly unclear—and no standard definition exists. The regulators rely on automotive companies to present a view of safety, while the companies themselves, each having a different interpretation of what constitutes

safe driving behavior, in turn seek input from the regulators. The majority of safety assessment today is self-reported by the testing companies, in good faith (Tesla Car Autopilot, 2018). These companies develop different interpretations of what constitutes safe driving behavior. Autonomous miles driven and miles per disengagements/intervention are two metrics closely watched by industry observers to provide a high-level view of AV safety. Interventions happen when either a safety operator detects bad behavior and takes control of an automated vehicle, or the vehicle itself detects something wrong and calls for a human to take over. Low rates of intervention do not necessarily indicate higher safety, they indicate only high agreement between drivers and automated systems. Humans can sometimes fail to detect hazards and if the automated vehicle fails too, they will agree without being safe. Therefore, disengagement is only an appropriate safety metric if the goal is to make AVs as safe (but not safer) than human drivers. It is no secret that safety assessments for automated vehicles need to evolve beyond the existing voluntary self-reporting. The hurdle is that there is no comprehensive common measuring stick to compare how far along each AV developer is in terms of safety.

Secnario2Vec

In this research, we first focus on capturing the temporal structures of traffic scenarios to form fixed-length vector representation. A scenario is defined as a short (< 10 sec) video clip captured from the front facing camera (plus possibility to extend to other sensing modalities). As indicated in Figure 1, the deep-learned visual features are extracted from the last fully connected layer from the CNN which takes image frames as inputs. In order to learn the temporal structure of videos, we utilize the Recurrent Neural Networks. RNNs are capable of preserving relations of each atom in learning sequential information since their specific structure of layers, which the current neuron outputs

depend on the previous computations. The learned vector representation is an encoding of the video clip. This encoding of a traffic scenario will allow us to "query" another AV data stream to extract similar scenarios and generate a standard of comparison between two AV software stacks and evaluate them on the basis of a standard safety metric.

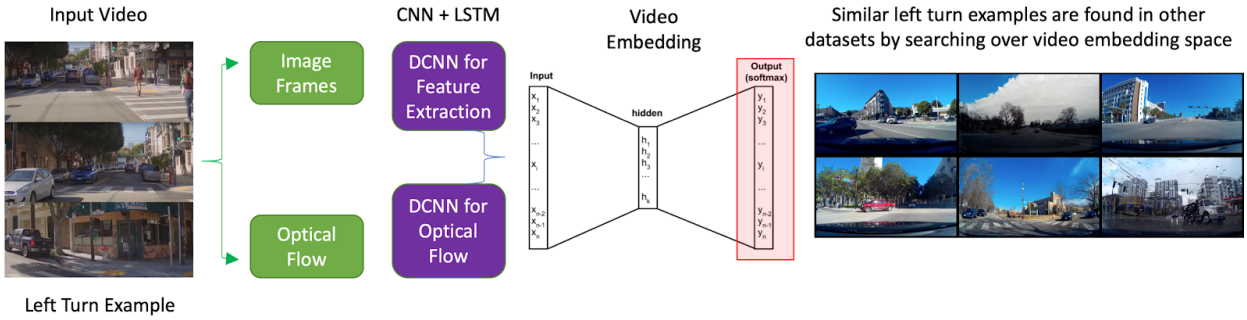

Figure 1. Architecture of suggested Scenario2Vec Neural Network

STS Project

The main purpose of this STS study concerning a standardized safety metric for autonomous vehicle software is to question the moral obligations and ethics behind developing a metric that will deem an autonomous vehicle 'safe,' and to examine and analyze the impact of a standard metric on the definition of progress in the autonomous vehicle field.

The first question will involve interviews with different actors in the autonomous vehicle field. This question is primarily aimed towards those who serve a primary role in the development of autonomous vehicle software. Currently, Waymo, GM Cruise, Argo AI, Tesla, Amazon, Apple, Uber and Lyft are among the few top autonomous vehicle companies that are developing and leading in this field. There are many more but for the scope of this research project, only representatives from the top companies will be chosen to interview. These companies will be representative of the 'commercial' sector of autonomous vehicle development.

Academic researchers also play a very large role in autonomous vehicle development. There are several professors at UVA, namely, Professor Madhur Behl, who specializes in

autonomous vehicles and robot research; he will also be interviewed. Lastly, consumers or potential users of the autonomous vehicle will also be interviewed as their definition of safety regarding an autonomous vehicle will be different. The interview will primarily focus on the individual's or companies' existing definition of 'safety,' and their current practices in order to ensure their vehicle is safe. The responses will then be analyzed for similarities and differences to determine where one software may be lacking in comparison to another and how the individuals incorporate ethical standards in developing their existing safety metrics. It is important to ensure that the individuals involved in the development of autonomous vehicles are developing their safety metric on not only the basis of 'efficiency' from a technical engineer's perspective, but are also incorporating moral principles to develop an ethical standard. The development of the safety metric informs the public of the autonomous vehicle's eligibility to be used and thus, the consumer is trusting the company/individual with their money, and their life. As a result, it is important to ensure the safety metric is representative of this cost.

Interviews for the first research question will also heavily focus on current and potential users of autonomous vehicles. Preferably, these users would be from the same companies that were interviewed to further highlight differences in the goals of actors in this research study. These interviews will involve asking users what they value and would look for when purchasing an autonomous vehicle and how a safety metric would impact their decision. These questions would also include how the user would go about buying a normal non-autonomous vehicle and if standardized safety metrics play a role in that (directly, indirectly, or subconsciously).

The second question will involve interviews with the companies mentioned previously, and academic researchers in the autonomous vehicle domain. The interviews for this question will focus on what the goals are of the individuals and companies developing the vehicles. This

will help show how the safety metric must be devised to consider these goals, and not ignore them. The introduction of a new standard safety metric can potentially change what the field is making progress towards. The standardization of the metric will inevitably introduce competition between leading autonomous vehicle manufacturers as each tries to improve the metric. As a result, it is important to determine what the metric is composed of as it will have an impact on how the field progresses. It is also important to define what this progress will be towards and what that will mean for the different actors in the autonomous vehicle domain. If the safety metric heavily weights one aspect of the car over another, it could have an impact on what developers choose to focus on, in a positive or a negative way.

This research study will rely upon two frameworks: "*Standards, Recipes for Reality*," by Lauren Busch (Busch, 2013), and "*Does Improved Technology Mean Progress?*" by Leo Marx (Marx, 1987). The first framework is going to focus on the first research question: "What are the moral and ethical obligations behind developing a metric that will deem autonomous vehicles safe?". This framework explains how standards are associated with power, and how all standards are of four types, and how standards are intertwined with ethics. This framework will be used in relation to a standardized safety metric for autonomous vehicles and to assess and evaluate the development of an ethical standard in this domain.

The second framework by Marx will focus on the second research question: "What will the standardized metric mean for the definition of progress in the autonomous vehicle field?". The framework that Marx lays out in his paper outlines two conceptions of progress. The first view during the Enlightenment era perceived science and technology to be in the service of liberation from political oppression. This view changed into scientific innovation being a basis for progress: "the social ends of new technology are not named, and we instead have a

"minimalist definition of civic obligation". Marx's framework presents a call to action for political, social and cultural goals comparable to those that were presented in the beginning of the industrial era to provide limits to the progressive vision of the future. He claims that although technology can mean progress, it is important to define what the progress is towards beyond efficiency, and what the purpose of the new technology is in order to progress socially.

The primary stakeholders associated with this technical project are developers of autonomous vehicles and potential users of these vehicles. This includes commercial manufacturers of autonomous vehicles such as Tesla, Uber, Lyft, and Waymo, and academic researchers in the field. The consumers will use the safety metric to determine whether or not to purchase the vehicle and will abide by their own internal standards to determine whether or not the vehicle is safe enough for them to use.

Conclusion

While the principle of safety by design (verification) is useful, it remains insufficient for automated driving systems, because of the existence of unknown scenarios that cannot be directly designed for, or verified. Therefore, the first goal of this technical project will be to effectively and efficiently generate driving scenarios that can be used to compare different autonomous vehicle software stacks. The second goal is to develop a standardized safety metric by which these autonomous vehicle software can be compared.

The primary goal of this STS research project is to determine which moral and ethical standards will need to be considered to develop a standard that will be socially accepted by the consumers. Furthermore, the frameworks mentioned previously will be used to investigate how the formulation of a standard safety metric will define what the field will make progress towards.

References

Tesla car that crashed and killed driver was running on Autopilot, firm says. (2018, March 31). *The Guardian.* Retrieved from https://www.theguardian.com/

Busch, L. (2013). *Standards: Recipes for reality*. Cambridge, Mass: MIT Press.

Marx, L. (1987). Does Improved Technology Mean Progress?". *Technology Review*, 33–41.

Wakabayashi, D. (2018, March 19). Self-driving Uber car kills pedestrian in Arizona, where robots roam. *The New York Times*. https://www.nytimes.com/