Proactive Planning and Control of Autonomous Vehicles
(Technical Paper)

The Introduction of Autonomous Vehicles into Society: Encoding Morals into the Machine
(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 in Systems Engineering

Katie Kleeman
Fall 2020

Technical Project Team Members:
Grace Glaubit, Noelle Law, and Jeremiah Thomas

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 _____ Date _____

Approved _____ Date _____
Nicola Bezzo, Department of Engineering Systems and Environment & Electrical and Computer Engineering

Approved _____ Date _____
Travis Elliot, Department of Engineering and Society

## Introduction

A technological trend that remains at the forefront of innovation these days is machine learning and artificial intelligence. More specifically, research and development of autonomous vehicles has been a large focus of academic, corporate, and governmental work. While prototypes have shown promising results, there still remains much work to be done for autonomous vehicles to have the capability to properly process the information collected, accurately assess the situation, and then react appropriately. This process is inherent to the human brain, and something we do quite well, however replicating this through software and algorithms is quite the challenge. These capabilities not only affect the efficiency of an autonomous vehicle, but also the safety - which is a main concern for both passenger vehicles and other robots. When using an autonomous robot, perhaps for exploration in dangerous areas or search and rescue, the costs resulting from a lost (unretrievable or broken) robot are largely monetary. However, when an autonomous vehicle is operating either within the presence of humans, or with human passengers, the cost of an unsafe journey can be far greater than simply losing hardware - it could mean injuries or deaths. For these reasons, the technical focus of this project aims to find a way to proactively plan and control autonomous vehicles to ensure a safe and successful mission, and the STS research project will investigate how we can program these vehicles to include morality and ethics in its decision making using the actor-network theory as a framework, as well as analyze how concerns about human safety will affect the successful (or not so successful) adoption of autonomous cars into society via the framework of social construction of technology (SCOT).

## Technical Topic

*Research Question*

As the development of autonomous vehicles has progressed, discussion of them becoming prevalent in society has increased. One major topic that is always associated with these discussions is the assurance of safety, even during unforeseen and unpredictable events. However, current technologies are not able to guarantee this safety in unknown terrains. In the presence of unforeseen changes in surface type, the autonomous system may need to switch and adapt its control strategy to achieve its desired goal and maintain a certain threshold of safety. For example, if an unmanned ground vehicle (UGV) planned a path that requires the traversal over a ramp, it should be able to both identify the ramp and speed up before it reaches it to successfully maneuver the surface. Furthermore, if the UGV is traversing over a low-friction surface (e.g. ice), it should identify that and adapt its linear velocity to go slower to keep a minimum traction as if it was on a high-friction terrain. In order to do this, the systems must be constrained at runtime to guarantee a safe operation that reaches its goal. Therefore, the research question we are asking is how can we proactively plan and control an autonomous vehicle in order to guarantee traversability and increase the chance of a successful and safe mission?

*Research Methods*

For the first semester, we will be using a virtual testing environment rather than testing on physical robots (that may change during the second semester depending on the status of COVID-19). The virtual environment we will be using is Gazebo, a three-dimensional robotics simulation. Scripts will be written in python and in tandem we will utilize the Robot Operating System (ROS), a middleware that sits on top of an operating system which is used to communicate with the robot. The first step is to run many experiments and collect data from

these experiments. These experiments will be run on a variety of different terrains with different obstacle configurations and at varying velocities. That data will then be used to train a machine learning model to recognize a certain external disturbance and adjust accordingly. The neural network will take inputs such as the pose and terrain type, and output the maximum safe velocity. Some tangible goals we have for this project include a learning-enabled framework for predictive planning and control, a simulator in Gazebo with realistic dynamics and environmental conditions for training and testing, and analysis in Matlab on training data and preliminary results. We have also been doing research into the relevant literature including Deep Reinforcement Learning for Safe Local Planning of a Ground Vehicle in Unknown Rough Terrain (Josef, 2020), Path Planning With Local Motion Estimations (Guzzi, 2020), and Terrain traversability analysis methods for unmanned ground vehicles: A survey (Papadakis, 2013).

## STS Topic

*Introduction*

As autonomous vehicle technology improves, it is important to analyze how this could influence, and be influenced by, society. The actor-network theory focuses on how different actors, human and non-human, interact with each other to create and affect technological development. Applying these tenets to this topic, my STS research project will investigate how we can program autonomous vehicles to include morality and ethics in its decision making, and how they will interact with other actors in the network. In addition, I will also be doing a Social Construction of Technology (SCOT) analysis on the topic. SCOT proposes the idea that the success (or failure) of technology is due to the social context the technology is developed, released, and used in. Using this perspective, I will research how concerns about human safety

will affect the successful adoption (or lack thereof) of autonomous cars into society and what other factors may affect this as well.

*Actor-Network Theory*

Actor-Network Theory (ANT) views "society" as an ongoing achievement, rather than a stagnant concept, and provides the tools to analyze the way in which society redesigns itself (Callon, 2001). It perceives scientific knowledge to be the effect of relationships and connections between objects, animals, ideologies, humans, social rules, and any other thing that affects technological development - these are the actors. Evolving the idea of constructivism in the 1980s, ANT's originating authors, Bruno Latour, Michel Callon, and John Law rejected the idea of social determination of scientific knowledge (Detel, 2001). The primary tenet of ANT is that the aforementioned actors interact with each other in a heterogeneous network. Moreso, the principle of generalized symmetry states that within that network human and non-human actors are to be given equal amounts of agency (L, 2007). The network operates via translations, when actors transform other actors, and by passing tokens (or "quasi-objects") between the actors. Tokens are created when the interactions within the network are successful ("Actor–network theory", 2020). When the actor-network is running smoothly, societal order is produced, but this order breaks down when certain actors are removed or tokens are failed to be transmitted (L, 2007). I plan to utilize this framework to analyze the introduction of autonomous vehicles into society. These vehicles serve as actors as they fit the following definition presented by W. Detel: "able to perform actions as a kind of behavior describable under some intention". Following the principle of generalized symmetry, I plan to acknowledge the vehicles as having equal agency as all other actors in this network including the law makers, potential users, pedestrians, and developers. Other non-human and ideological actors include morals and ethics, safety, social

rules, and influential institutions. In my paper I will investigate the most prominent actors involved in this network and how they might translate each other, pass tokens, or generally interact to contribute to the development of autonomous technology.

*Social Construction of Technology*

Drawing on Actor-Network Theory and inspired by the principle of symmetry from Sociology of Scientific Knowledge, the Social Construction of Technology framework holds that the success of an innovation is not because it "works" better than failed innovations, but rather is due to the social context that promotes (or fails to promote) it ("Social construction of technology (SCOT)"). To that point, SCOT aims to investigate and analyze the metrics used to judge technology, which groups or stakeholders defined them, who is included or excluded from that group, and why they defined it in a particular way. It proposes that human action shapes technology, and its uses need to be understood in the social context. The leading scholars on this topic, Wiebe Bijker and Trevor Pinch, began this work as a response to technological determinism, the idea that a society's technology determines its social structure and values. Disagreeing with that viewpoint, they adapted the Empirical Programme of Relativism (EPOR), a method that analyzes how scientific findings are socially constructed, and produced the SCOT framework (Klett, 2018). A few key tenets to this concept are relative social groups, interpretive flexibility, and stabilization. Relative social groups are defined as "all members of a certain social group who share the same set of meanings, attached to a specific artifact" ("Social construction of technology (SCOT)"). This is important in my analysis as many different social groups will be involved or affected by the introduction of autonomous vehicles into society. These different groups may attach different meanings and morals to the problem at hand, which leads to the next tenet. Interpretive flexibility states that there is no ground truth "best" artifact,

technology, or methodology, but rather each group defines what they mean by "best" based on the interpretation of the problem they are trying to solve. I believe this will be important in my research as different groups of people from different cultures may have different definitions of morality or ethics, and may also have different views on the purpose of autonomous vehicles and how beneficial they could be to society. Finally, stabilization occurs when one social group dominates the others, and therefore their ideals and design prevails while the others are forgotten ("Social construction of technology (SCOT)").

*Research Plan*

For this project I intend to investigate and analyze the introduction of autonomous vehicles into society, and the social, technical, and ethical factors that contribute to (or inhibit) their success. In particular, I hope to research how different cultures and groups of people define their moral guidelines and ethics. A research study called "The Moral Machine" is a survey of 2.3 million people that reveals there may not be a universal moral code ("Moral Machine"). Using this as a basis for understanding, I will see how these values are encoded into the algorithms of autonomous vehicles and will evaluate that and its effects using the frameworks discussed above. Amy Maxmen's article, "Self-driving car dilemmas reveal that moral choices are not universal", mentions that self-driving cars are projected to cause fewer accidents, proportionally, than human drivers. However, due to the disproportionate media coverage, people may not trust the safety level of autonomous vehicles (Maxmen, 2018). Viewing this using the SCOT framework, one could guess that these social factors may hinder acceptance of this new technology. As autonomous vehicles have developed more over the years, society gets closer and closer to actually integrating autonomous vehicles into regular traffic. However, with multiple deaths and injuries arising during testing, it is crucial that we understand how these cars

are programmed to react and which values they will be prioritizing. From this research, I hope to explore the following key questions: How do we define our morals/ethics as a society? Technically speaking, how do we code those morals into the machine learning algorithm? How will these discussions on safety affect the success of autonomous vehicles in society? Some sources I intend to use include "The folly of trolleys: Ethical challenges and autonomous vehicles" by Heather Roff, "The everyday ethical challenges of self-driving cars" by Johannes Himmelreich, and "Machine Ethics: The robot's dilemma" by Boer Deng.

## Conclusion

The focus of my technical project is proactive planning and control of autonomous vehicles, towards the purpose of increasing and ensuring safety during operation. Linking back to my STS project, the concern of safety is paramount and crucial to the acceptance of autonomous vehicles in society. Therefore, if our research is successful in ensuring a certain threshold of safety, it and other works could lead to a better public perception of autonomous vehicles and their uses in society. Using my technical knowledge of the problem as a foundation, I hope to use my STS research to analyze how morals and ethics can be encoded into autonomous vehicles, how these algorithms and vehicles serve as actors in a greater network, and how people's perception can contribute to the success or failure of self-driving cars in society.

## References


Actor–network theory. (2020, October 09). Retrieved October 18, 2020, from https://en.wikipedia.org/wiki/Actor–network_theory

Callon, M. (2001). Actor Network Theory. *International Encyclopedia of the Social & Behavioral Sciences,* 62-66. doi:10.1016/b0-08-043076-7/03168-5

Deng, B. (2015, July 1). Machine ethics: The robot's dilemma. Retrieved October 18, 2020, from https://www.nature.com/news/machine-ethics-the-robot-s-dilemma-1.17881

Detel, W. (2001). Social Constructivism. *International Encyclopedia of the Social & Behavioral Sciences,* 14264-14267. doi:10.1016/b0-08-043076-7/01086-x

Guzzi, J., Chavez-Garcia, R. O., Nava, M., Gambardella, L. M., & Giusti, A. (2020). Path Planning With Local Motion Estimations. *IEEE Robotics and Automation Letters, 5*(2), 2586-2593. doi:10.1109/lra.2020.2972849

Himmelreich, J. (2020, April 16). The everyday ethical challenges of self-driving cars. Retrieved October 18, 2020, from https://theconversation.com/the-everyday-ethical-challenges-of-self-driving-cars-92710

Josef, S., & Degani, A. (2020). Deep Reinforcement Learning for Safe Local Planning of a Ground Vehicle in Unknown Rough Terrain. *IEEE Robotics and Automation Letters, 5*(4), 6748-6755. doi:10.1109/lra.2020.3011912

Klett, J. (2018, July 20). SCOT. Retrieved October 18, 2020, from https://stsinfrastructures.org/content/scot

L, D. (2007, March 23). Actor-Network Theory (ANT). Retrieved October 18, 2020, from https://www.learning-theories.com/actor-network-theory-ant.html

Maxmen, A. (2018, October 24). Self-driving car dilemmas reveal that moral choices are not universal. Retrieved October 18, 2020, from https://www.nature.com/articles/d41586-018-07135-0

Moral Machine. (n.d.). Retrieved October 18, 2020, from https://www.moralmachine.net/

Papadakis, P. (2013). Terrain traversability analysis methods for unmanned ground vehicles: A survey. *Engineering Applications of Artificial Intelligence, 26*(4), 1373-1385. doi:10.1016/j.engappai.2013.01.006

Roff, H. M. (2019, October 25). The folly of trolleys: Ethical challenges and autonomous vehicles. Retrieved October 18, 2020, from https://www.brookings.edu/research/the-folly-of-trolleys-ethical-challenges-and-autonomous-vehicles/

Ryan, C., Murphy, F., & Mullins, M. (2020). Spatial risk modelling of behavioural hotspots: Risk-aware path planning for autonomous vehicles. *Transportation Research Part A: Policy and Practice, 134*, 152-163. doi:10.1016/j.tra.2020.01.024

Social construction of technology (SCOT). (n.d.). Retrieved October 18, 2020, from https://web.archive.org/web/20180410205247/http://www.stswiki.org/index.php?title=Social_construction_of_technology_(SCOT)