Approaches to Define Ethical Principles in Self Driving Cars?

University of Waterloo

Chantelle Gellert
Dr. Daniel M. Berry

I Introduction

Part of growing up is becoming more independent, having increased responsibilities and, in turn, acquiring extra freedoms. One of these freedoms is the option to drive somewhere rather than utilizing public transportation, or relying on the generosity of parents, friends or significant others. The ability to drive provides significant flexibility, not just in one's personal life, but also opens up new opportunities for employment. As this flexibility is often seen as an expected part of adult life, driving has become almost a basic necessity for an adult in modern society. With the use of automation and machine learning to replace or make day-to-day tasks easier, it comes as no surprise that automatic cars are on the to-do list for enhanced driving experiences. As artificial intelligence (AI) becomes a prominent feature for self-driving cars, questions relating to the morality and ethics of the artificial agents become more relevant. (Gellert, 2017)

It is a difficult task to develop a thinking machine that can understand and obey the laws for driving on a road. When exploring the research area regarding how self-driving cars will interact with human life, one should think about the ethically derived practices which will be encoded into the machine. (Pennachin, 2007) Although ongoing research is being performed to fully understand how best to tackle the ethical dilemmas involved with driving, this report will explore different issues that arise when teaching machines ethics. As ethical beliefs typically differ between cultures and individuals, it would logically follow that ethics for humans would differ from ethics for machines. As humans already have a generally accepted set of driving standards, it is reasonable to use these principles as a suitable baseline for a machine operated vehicle. Since human ethics are still not a complete set, it is difficult to create a defined set of rules. Nevertheless, an ethical decision must be made given extremely limited time constraints. (Neurosci, 2017) Machines, however, are seen to be more capable of making the most optimal ethical decision as they are able to appraise a vast number of alternative solutions to an ethical dilemma in a shorter period of time when compared to a human. (Hurtado, 2016) In the event of an accident, error, or malfunction, who will be liable for the ethical decisions made by the AI, and how will the AI's ethical decisions aline with those of human drivers?

In the remainder of this paper, Section II defines ethics and why it can be arduous to translate ethics into machine rules. Section III segregates the challenges and liabilities with self-driving cars. Section IV proposes potential solutions to the challenges identified in Section III. Section V briefly highlights further exploration ideas and Section VI presents the results.

II Ethics

As seen in stock market exchanges, AI systems are able to make critical decisions that impact the environment around us. The capabilities of AI systems will only increase, which raises the importance of embedding ethics, such as "All human life is precious," into the systems. Throughout life, it is generally accepted to let our moral compass guide our decisions and our moral compass is guided by our ethics. Although morality and ethics are often used interchangeably, they do have different meanings. Morals are what is considered right, wrong, good, or evil to an individual. Morals are unique to a person as they are developed by the individual's personal experiences throughout life. (Weiss, 1942) Alternatively, ethics is a systematic review of values or a subset of morals which a society agrees upon. (Shuriye, 2012) One would be considered to be acting ethically if they conform to those values set the society. (Thomson, 2001) In a more general sense, there are two main views of ethics in the world, eastern views and western views. (Shuriye, 2012) Eastern

views are based on many religions such as Buddhism, Confucianism, Hinduism, and Taoism. Whereas, western views are mostly based on Judeo-Christian philosophy. Thus, it is important to note the self-driving cars imported from another country will have foundational differences making it even more difficult to address ethical concerns in software. Therefore, it becomes difficult to decide upon a single ethical view that is globally applicable.
There are two important viewpoints to consider with ethical programming: how ethical responsibility is applicable to the system governing self-driving cars, and the challenges of translating ethics into something a machine can understand and execute.

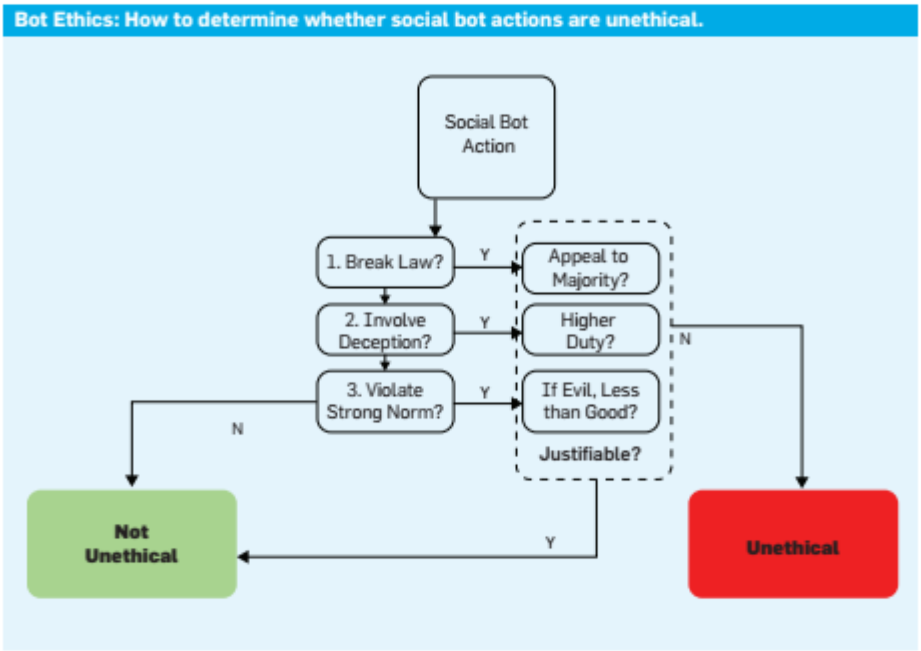

Figure 2.1: Bot ethics (Lima Salge, 2017)

As artificial intelligence advances and becomes sentient, it can develop its own set of ethics that could be different from human ethics. However, since humans understand their own set of ethics, humans will model AI systems based on them. As seen in figure 2.1, a bot ethic system is presented in order to show how an AI system can handle programmed ethics. In order to have an ethical evaluation take place, an AI needs the ability to reflect and debate upon an action before performing it. The most basic ethical question: “Does the considered action break a law?” However, if the bot performs an illegal act, then the AI is required to justify the action in order to adhere to the moral standard of the majority. In turn, there are ethical dilemmas that do not necessarily break the law. An AI can violate the ethical rules of being trustworthy or truthful through deception. However, deceitfulness can be justified if the act is superseded by a high-order ethical duty. For example, an AI could exceed the speed limit to avoid a collision from the rear. Lastly, can an AI’s action be performed both legally and truthfully, and still be considered unethical if this action violates a strong social norm? For example, Tray is a social bot created by Microsoft who started tweeting statements such as “Hitler was right I hate the Jews.” Although Tray believed this to be true and had the right to free speech, people look down upon certain statements and Tray was removed from Twitter. Thus, concluding that Tray’s actions were indeed unethical as there has been the age old debate about just vs. unjust, good vs. evil. In the end, it's mostly determined by the majority. (Lima Salge, 2017) (Kraft Amy, 2016)

Ethics is a complex domain to deal with fully, but there are some ethical behaviors that are unacceptable to be compromised, such as sacrificing a human life. Even so, blocks can be found in ethical programs. Finding more understandable logic for the AI to follow has proven difficult while learning is still commencing.

III Challenges and Liabilities

iii.i A perfect world

In a perfect world, we would not have to code ethical dilemmas, since we would never encounter the case in which we had to sacrifice one person's life for the sake of another. In theory, it is possible that if all self-driving cars were to drive in unison, communicate perfectly with each other, and follow a set of rules, no collisions would occur. Due to the nature of self-driving vehicles predictability, one does not have to worry about self-driving cars having the erratic behavior of humans such as reckless driving (speeding, failing to stop at a red), impaired driving (drowsiness, drunken state), and failing to correctly interpret the intentions of the other vehicle. In order to construct the perfect world of zero collisions, the external factors also need to be in a constant state such that self-driving cars will not be affected by anomalies such as by weather. Due to extreme weather, including floods, tornadoes, and blizzards, it may be impossible for the car to follow the set of instructions it receives. For the perfect world of self-driving cars, one would also have to control the behavior of the citizens or be able to predict the citizens' intentions. Unlike self-driving cars, humans are born with free will and their foundational core values vary greatly from one another. It is a foolish assumption to think that all humans will follow the rules set out for pedestrians, as it is clear that pedestrians are found on roads every day for all types of reasons, from jaywalkers to children playing games on the road, including chalk coloring or a more common game in Canada, road hockey. (Craggs Samantha, 2017) Needless to say, self-driving cars cannot accurately predict human behavior because no one can say for certain what anyone will do without future knowledge. Even with continuous trajectory predictions, a human can change course for any number of reasons. Thus, the erratic interactions create a real problem that can stir up ethical debates about how to deal with these human interactions in an ethical manner.

iii.ii Pedestrians don't obey law

An example of a self-driving car interacting with an erratic pedestrian is given in the case of a woman who crossed a road with two lanes of traffic outside of a crosswalk in the dark and was struck by a self-driving car. (Levin Sam, Wong Julia, 2018) We could say that self-driving car should be able to recognize pedestrians. However, how do we judge the self-driving car for failing to slow down in the case where a pedestrian proceeded slowly across a 40 mile per hour road at night while wearing dark clothing? We could say that, in the perfect world, the LIDAR (Light Detection and Ranging) laser sensor which can be used in the dark would be able to detect an object impeding its path and stop since it is programmed to stop for pedestrian crossing or bicycles. (Marshall Aarian, 2018) However, is it really the self-driving car's fault for failing to register the pedestrian, or should we fault the pedestrian for crossing illegally, the Uber driver for not being properly alert in order to respond, or Uber itself for not properly training or monitoring the behavior of their employees?

iii.iii Recognition problems

Currently, the software is not perfect at detecting humans and people can be quick to blame the faulty software, but the problem is non-trivial. For example, the human could have crawled across the road on all four limbs like an animal, such as witnessed in a personal experience where a drunken person attempted to crawl across the road, outside of the cross walk, into oncoming traffic. The human was acting like an animal, but other humans would still be able to recognize them as a human. For another example, many people dress up as another entity such as a mascot or wear a costume during Halloween. Halloween is a common tradition in North America when children dress up as all sorts of creatures, superheroes, or monsters for fun. In general, the car should be able to detect any object in its way, as well as any object following a predictable route which would collide, and thus execute an appropriate action, if it can be performed safely, such as, slowing down, changing lanes to avoid the object, or stopping at the side of the road and waiting for assistance on how to proceed.

iii.iv Trolley Problem

Some of the largest issues with ethical dilemmas are who to blame when someone dies as well as computing the value of a human life, since there will eventually be a situation where someone's life has to be taken. For example, in the famous trolley problem, should one divert the trolley and kill one person, or do nothing and have the trolley kill five people? When relating the trolley problem to self driving cars, the self driving car might have to decide who will die between occupants of the car and pedestrians.

Figure: 3.1 Japan, Kiyotaki Tunnel (Sagakiyotaki-ku Kyoto-ken, 2017)

Imagine a poorly lit, single lane tunnel where the vehicle has very little room on either side. It would be extremely unsafe to enter this area as a pedestrian. However, even with many signs urging caution, there could be a person walking in the middle of the road. Due to the poorly lit nature of the tunnel, the self-driving car can only see so far ahead (as seen in the Arizona crash when the LIDAR was not able to identify the woman in the dark lighting) and it might be faced with a choice between continuing toward the pedestrian or crashing against the wall which might kill the occupant of the car. Despite proper warnings signs and common sense dictating not to enter, this case still will be debated to determine who was at fault, the pedestrian for entering the tunnel, the car for not recognizing the danger in time, the tunnel creators for not having better lighting, or the warnings to deter pedestrians. Everyone and everything can and will be accused of fault. Whoever is chosen to die, their family and friends will be upset since they do not see it as the program making a difficult choice in a situation, but rather the program making a deliberate choice to kill their loved one. Instead of comparing the program to an imperfect human that made their best decision, the family and friends would expect higher standards. Why does a program need higher standards than a human? Is this due to lack of control on the humans part? However, everyone made a choice that resulted in the accident. The pedestrian entered a dangerous area which is also accessible to self-driving cars, the occupant of the car conceded the ethical decision to the self-driving car, and the company constructed the tunnel with poor lighting. Should everyone share part of the blame?

iii.v Ethical issues with trains can foreshadow self-driving car

The ethical debate is not limited to self-driving cars since we find the same issue when dealing with trains, where it is not the pedestrian's fault even when they did not obey safety precautions. When a subway car is coming into the subway station and a person falls in front, there is sometimes no possible way for the train to stop in time depending on when the person was noticed on the tracks. Additionally, there is also nowhere for the train to go except to follow the rails. For example, (Nir Sarah Maslin, 2017), a woman who drove onto train tracks, required a two year long investigation to finally be found at fault in a fatal Metro-North crash according to investigators. Even though the driver of the vehicle was spotted by eyewitnesses exiting her vehicle and then continuing forward onto the train tracks, it was not a simple case to assume the driver of the vehicle was to blame for the accident. Even though the official investigators of the case concluded after two years it was the driver's fault, the driver's family still conducted their own investigation to blame the Metro-North railway for her death. When pedestrians or cars unlawfully enter the train's right of way and the train could not possibly stop in time, it is not assumed that the railway will be exonerated from any fault. Even though it is a criminal offense to trespass on railroad tracks, railways are not automatically exonerated from any incidents that occur and each incident needs to be heavily investigated. The law also imposes burdens on train operators, including the duty to: keep a proper lookout for obstructions on the track and take reasonable steps to avoid striking a person or object on the tracks. (Grossman Michael, 2015) People tend to improperly cross railroad tracks or roads outside of designated areas when they do not see any oncoming traffic. Since these people improperly cross so frequently without consequences, they become complacent and ignore the designating crossing areas which puts them in jeopardy, forcing self-driving cars to need to take unpredictable pedestrians into consideration. (Orf Darren, 2014)

iii.vi Mapping perfect ethic system

It is expected that eventually having to make a choice over who to kill is currently inevitable and thus the ethical decision of which person the self-driving car decides on should be predetermined. However, if the decision over who to kill is already predetermined by the program, does that mean the programmer is responsible for the death? Or should the supplemental operator become responsible for agreeing to a complex rule set? Do the pedestrians have a say what ethical choices are predetermined in the car? It is unreasonable to assume that the supplemental operator will fully understand the car's ethical system if it is not the same ethical system as their own, let alone agree with it. A proposed solution could be to program a set of ethical rules which directly correspond to the supplemental operator's unique view point. However, self-driving cars can have many different users in addition to the owner. Also, it is important to keep in mind that our ethical systems evolve as we gain more knowledge about ourselves and the world. Therefore, to truly match each supplement operator's unique viewpoint, a new ethical scan will need to be done each time the car is taken out for a drive. It is important to note that the car will still drive following the rules of the road, but when faced with an ethical dilemma outside of the rule set, it will refer to the modeled brain of the supplemental operator for a decision. However, even if the accident actually occurs based on a person's exact ethical modeling, it is reasonable to believe that the person will deny fault for the person killed when they were not physically driving the car, even if we can prove the action the car took was from perfect copy of the supplement operator's ethical system. In a fictional movie called "Minority Report", there is a crime stopping team that can predict the future and arrest people before crimes are actually committed. In this fictional universe, although crime rates went down, how can one prove that they would have actually committed the crime unless they physically go through with it? We can extrapolate this to support that encoding the proper ethics into the car is not a perfect solution to the debate of who is at fault for accidents when the supplemental operator relinquishes control to the car itself.

iii.vii Reasonable age to interact with a self-driving car

Consider the case that the new ethical scan system can be done with a young person who has not fully developed. At what age will you be able to operate a self-driving car? It is safe to assume a child is not allowed to drive a car because they have underdeveloped fine motor skills which will diminish their reaction time in an event of an accident. Furthermore, a child starts to develop critical thinking skills and moral decision making between the ages of 11-15 (Kozier, Erb, 2014). With self-driving cars, there is no longer the same intellectual or physical requirement but ethical requirements remain. If a young person is the supplemental operator of the self-driving car and an accident occurred resulting from an ethical choice, is it now the responsibility of the car, the supplemental operator, or the legal guardian? The minimum age required to start driving legally in North America is typically sixteen (fourteen in Alberta, Canada), but the age to make an ethical decision about the direction of the country (voting) cannot be taken until eighteen. Self-driving cars would be equipped with the knowledge and capabilities to drive in ways most humans can only imagine, and now it could be guided by the ethical mindset of someone the government has deemed not qualified to make higher levels of ethical decisions like voting (lowest allowed age is eighteen across both Canada and United States), the use of tobacco (lowest age allowed in Canada or the United States is eighteen), and alcohol use (lowest allowed age is eighteen in Canada and United States sets the minimum legal drinking age to 21 for all states). (USA Government, 2018) (Tobacco21, 2018) (NIH, 2018) (André Blais 2001) (Royal Canadian Mounted Police, 2016) (LCBO, 2018) (Government of Canada, 2017) As a direct comparison, the enrolment age in the armed forces without parental consent in North America is 18. Military personnel frequently make ethical decisions regarding life and death, such as the self driving car could referencing back to the trolley problem. (Service Registration, 2018) (Government of Canada, 2017) With the misalignment between age and ethical decision making, will there have to be an ethical standard to comply to? Will the age of driving have to be raised or will a new driving test need be passed before being allowed to operate a self-driving vehicle? How the government would measure ethical standards and would having a set of "guideline" ethical standards be considered brain washing?

iii.viii Self-driving car is for people that can't drive

People who can not drive still need to go places, such as doctors, grocery, visitations and other activities outside of the home. It can be a non trivial task transporting people that can not drive on their own around. Self driving cars can become a solution for those with physical or mental disabilities to travel however, where will the ethical responsibility lay when the occupant of the vehicle is not capable of reaching the bar of minimal ethical decision making to be a occupant of the self driving car? Even if you can not physically take over for the self driving car but meet the minimal ethical requirements can you still be a sole occupant of the self driving car?

iii.ix Prejudice

One would think that the car would have the same ethical values as the owner even if the owner is not the operator of the vehicle at all times, especially if the owner is responsible for all activities of the vehicle. However, if the vehicle mirrored the owner's ethical standpoint, would it be prejudiced towards certain people? Racism is a factor to consider when a person is faced with an ethical debate about whether to kill someone from race A or race B. Even if racism isn't a factor, will people still be accused of it when choosing self-preservation over killing someone of another race? What about other forms of prejudice including but not limited to, sexism and ageism? Who's ethical rights trumps the others?

iii.x Self driving cars can bring a net good

Even with the complications caused by mapping one's personal ethical values or by using the self-driving car's recognition software, measures could be taken to advocate the benefits of a self-driving car rather than focusing on the cons, since an acceptable ethical threshold would inevitably increase the overall safety of the vehicle. We can implement an accepted system of ethics for self-driving cars and evolve it based upon events. An ideal set of self-driving car ethics cannot be determined without testing which will be difficult since there is a large enough controversy for animal testing without including human testing. Reportedly 25,000 virtual self-driving cars are making their way through fully modeled versions of Austin, Mountain View, and Phoenix, as well as special test-track scenarios for particularly tricky roads. (Madrigal Alexis, 2017) However, these are all simulations where no harm can be done towards humans and result in no practical experience with the unpredictability of the real world. It is important to note that there are people who already risk their lives for testing technological advances, such as in new medical treatments to increase the odds of survival. Self-driving cars have a baseline of consistent driving standards which will be higher than some of the worst human drivers who are still allowed to drive on the road. If that threshold is not met, then it is obvious that the cars should not be allowed on the roads.However, if the threshold is met, one can assume that the technology has the qualifications to drive like other humans. Self-driving cars could enforce a new level of driving standard which will keep the roads safer by replacing drivers who fall below the self-driving car standard for driving, which will increase the odds of survival. A higher standard of driving will hopefully increase the overall safety of our roads, especially for people who suffer from conditions preventing them from driving optimally, such as intoxication or disability. In roundabouts, some people are more aggressive or don't pay attention, but if you have self-driving cars communicating with each other you will know what exit a person is about to take so you know when it is safe to enter. A lot of the time with roundabouts, cars are in the center lane and you do not know if they will continue around or be exiting straight and this causes confusion, especially when people do not use signals properly. Even though roundabouts reduce the amount of fatal accidents due to lower speeds, they generate more overall collisions. (Metcalfe John, 2017) Self-driving cars will also help with distracted drivers. Today, most people stop and check their phones when they are at a red light or are in a hurry to get some place because they have tasks to complete. (Latimer Kendall, 2016) Self driving cars could allow the supplemental operator to safely interact with their phone or complete other tasks while traveling. Since the supplemental operator no longer needs to concentrate on driving and can advert their attention other tasks, self-driving cars will reduce the number of people ignoring safety laws. Even though you cannot realistically get a world of entirely self-driving cars because people enjoy driving cars as a sport and pastime, an increased number of more safe vehicles on the road should save lives which is the main focus of the self-driving car methodology of getting from one place to another safely.

Furthermore, taxis are often expensive, but often the only option for those who can't drive a car. Self-driving cars will bring more accessible modes of transportation for those incapable of driving, such as when they are intoxicated or have a permanent/temporary disability. Hopefully, encouraging humans to take safer methods of transportation around town.

iii.xi Weights of life

Another interesting perspective is weighing the worth of everything involved in an accident. How do we compare humans, animals, and objects? Of course people would want to live if given a choice, but what if someone else had to die in order for you to live? Should there be a choice to preserve one's own life above all others? Using the example before with the tunnel, the car can choose to kill the pedestrian on the road or the occupants of the car, but what if the child of the supplemental operator was in the passenger's seat? Even though one would think most people will choose self-preservation over the survival of others, this is not always true since it is fair to say that a parent might sacrifice themself to save the life of their child. For example, the car could have a choice to run into the wall on the left and kill the driver or into the wall on the right and kill the driver's child. Will the car need to know all passengers priority value and seating location? What if the car crashed on the child's side and the odds of surviving the crash were 90% for the parent and 0% for the child compared to if the car crashed on the parent's side and the odds of survival were 70% for the child and 0% for the parent? Should the car respect the priority system or choose the greater chance of preserving life? Should there be a ratio? Also, it is possible to change seat positions. When will the car calculate priority of passengers? If it's possible to change seats at any time, should the car constantly monitor seat positions? Will a debate have to occur about which person is more important and at what survival rate should their life be considered more important than another person, with each seat positions change? If a rule set has to be agreed upon, what should the car do when it notices a change? After a rule set has been input, how can one guarantee that all parties were in agreement to the settings? In addition, similar rules will have to be set for the car when choosing to kill pedestrians or passengers. It is the same situation as before with the parent and the child, but a child can chase a ball into oncoming traffic and the car has to decide to choose a passenger or the child. Younger people have longer to live, but older people could have a valuable skill set or knowledge which needs to be protected. Who will take precedence over the other and will there be setting which has a special encoded list to protect certain individuals above everyone else, like the president of the United States? Taking a look at the situation further, when the value of one individual is to be above all others, how far will that go? Allow 5 deaths to save the president? 10, 50, 100? How valuable will one person life be considered and how will this evaluation be made? Will people be given a lower score based on a health condition? Such a scoring system is bound to start to affect other areas of your life such as life insurance because every self-driving car in the world might place you at a lower priority compared to everyone else. It might devolve into an absolute ranking system and people would really want to know how they rank amongst the total population in the world.

Also, don't forget to program your self-driving car to protect your family cat, Mittens. Most people will not sacrifice themselves for an animal, however some people may think their life is less important than animal. Some religions or personal beliefs may consider certain animals as more important than human life in general. Ethical debates will happen if you are not allowed to program the car to protect animals, but what will happen if you kill your passenger to save Mittens the cat? Would you liable for your passenger's death, or would the program be, simply for following instructions that you set? In addition, the passenger may not have known that they have been signed up to be sacrificed for an animal and died because the ethical rules set for the passenger were not checked first before they entered the vehicle. This could cause people to fear carpooling and increase the number of single rider vehicles. Going back to the Halloween example, a child could be dressed up as a cat and the car could be programmed to ignore animals and which might result in the car failing to stop for the child. In this example, does the responsibility for the child's death rest with the car for not recognizing the "cat" as a human dressed up or the driver for not wanting the car to stop for animals? The Endangered Species Act makes a select set of animals illegal to kill, harm or otherwise "take," so who becomes responsible for not following these guidelines when an endangered species is harmed or killed? (Gill Kathy, 2017) Typically, programmable electronic devices come with a base set of settings, better known as the factory settings. For a set of ethical rules, if the default setting is never changed and the car makes an ethical choice in a way that the driver did not intend, who is at fault? The driver for not setting their car properly or the default settings set by the maker of the car?

IV Imperfect solution

Even with all possible knowledge, there will still be inquiries about ethical issues related to self-driving cars, which begs the question: how will self-driving cars ever be able to move forward towards becoming universally adopted?

iv.i Self driving cars are better but not perfect

The best way to solve the ethical dilemma of self-driving cars is to have open-source resources to help the general public understand the program. Self-driving cars are not and never will be perfect as the navigation software is written by people who are not perfect. In addition, there are infinitely many different scenarios that cannot realistically be accounted for and thus adopting a no fault insurance system is proposed. Since deaths and accidents in general should decrease by an astronomical amount if self-driving cars became universally adopted, then, in theory, insurance companies would be saving money overall and should just pay for any damages that occurred no matter what happened. (Lafrance Adrienne, 2015) There could be premiums based on the car's make and model based upon its software record instead of an individual's driving record as they are not driving. Assuming the occupant never has to take over driving, the insurance company is not insuring a person, but rather a software package, the rate should, in theory, be better for most individuals and there should be no risk of insurance rates increasing which should be more appealing to people. In theory, the occupant could never have to drive since certified individuals sent by insurance companies can take over or specialists can come tow the car. You are not at risk of being at fault unless you maliciously attack the software, and even if it was attacked by someone else, such as hackers, your insurance won't increase. Everyone would pay the same amount of money, which will more evenly distribute the cost of insurances. Ideally, the cost will go down for everyone, as less accidents means less insurance money used for claims. The insurance is on the car which will cover the case if there is no passenger at all. Ideally, the insurance company should be regulated and consistent with paying out the insurance money. It will take significant financial backing, but the projected result could unify the insurance framework to aid when ethical decision making goes wrong. Another factor to consider is whether or not the proper updates for the software or regular maintenance was performed on the car. Given that the car is a self-driving car, if the maintenance is not performed a reasonable amount of time before the car becomes too dangerous to drive, then the car can drive itself to an update shop where it will get the updates and maintenance needed. If you cancel the updates or refuse to perform the required maintenance, then you are voiding your eligibility for the no fault insurance just like you can currently void a warranty until the car is properly inspected, it may even be best to disabled self-driving mode at this point. Software updates should in theory be automatic as cars can be equipped with Wi-Fi capabilities in order to allow update patches or link to a satellite in areas where Wi-Fi connections are not established. It is important to note insurances will need to protect against malicious attacks (hacking).

iv.ii Self preservation

One concept to indulge is to give the cars the idea of self-preservation which has been debated as part of human nature. If the car attempts self-preservation, then there is no bias towards who the car has to choose to kill since the car would choose the path of least damage to itself. By avoiding damage to exterior objects such as pedestrians, we hope to minimize the damage to passengers as well. If the car does not take much damage, then there is less chance of harm to the passengers. The car did not choose to inflict harm to its passengers nor did it choose to harm the pedestrians, because the car made a selfish but unbiased decision to mitigate damage to itself. Ideally, no one should ever be sacrificed for another, but it is unrealistic to ignore the facts of it happening and it is unreasonable to leave it up to a random generator to decide.

iv.iii Privacy vs Safety

The Chinese have built Facial Recognition software to fine jaywalkers via text message with the intention of cracking down on pedestrians who are acting unlawfully. Although it is meant to increase safety, it is also touching on the earlier notion of safety vs. privacy, as the Chinese are using extensive surveillance networks to track the pedestrians and invading privacy by messaging personal numbers with the fine. (Grossman David, 2018) Which leads to one possible solution to solve the non-trivial recognition problem is to implant chips into all humans for the self-driving cars to pick up the signals from to determine if the object it detects is a human and react appropriately. However, this will cause a whole new ethical dilemma and debate about privacy and most people will not volunteer to be constantly tracked and monitored, even if it provides a safer environment when dealing with self-driving cars. However, in general privacy from the government does not truly exist. It is important malicious attacks (hacking) will cause safety risks to occupants of the vehicle and pedestrians.

iv.iv No fault does not mean no justice

The no fault system does not refrain from assigning blame as there needs to be a sense of justice as well as the proper procedures followed such that the incident does not happen again. The no fault system ensure that everyone is still insured at the same rate in the event that an incident occurs as long as there was no malicious intent (Insurance fraud). When determining blame, the same principles of courts today are applied, given the circumstances, what decisions were made against the existing laws, who is responsible for the incident. In a similar way to existing court trials that make ethical decisions about what would be wrong given the presented information, the no fault system for self-driving cars would open investigations and pay compensation appropriately. The idea of no fault insurance is that everyone shares a part in the self-driving car mistake as everyone shares part in them being allowed on the streets. Paying insurance premiums should not only fall to the drivers, but also the manufacturers of the car as they bear some responsibility from building the system and need to back a guarantee of minimal functionality. For example, if the car was built without airbags the manufacturer bears the responsibility for not providing a product up to a minimal standard.

iv.v Misinterpreted rules

Reassurance of compensation paid hassle free for mistakes by self-driving cars and humans will hopefully increase acceptance of self-driving cars on the road. A no fault policy in effect should not assign blame, but focus on rectifying the situation, which will ideally shift the attitude towards the situation and not the people. For example, a typical rule for a road could be a speed limit of 40 miles per hour, but when construction is taking place the road speed limit rule could be 20 miles per hour temporarily. Another rule change can be found with a crossing guard holding a stop sign in a nonstandard crosswalk for pedestrians to cross and temporarily interesting the rule to stop to oncoming traffic. The car will not always recognize these changes to the road instructions, but humans may not correctly interpret or notice these changes either. As a result of the misinterpretation, someone might get hit by the car even though the passenger expected the car to follow the law for the temporary road instructions. The no fault will hopefully bring about an equilibrium between those involved in the situation.

iv.vi Postponing decisions

A great way to deal with ethical debates regarding AI decisions (including recognition problems) is to postpone the situation and wait for assistance. During a situation in which it is impossible for the self-driving car to proceed forward on a given path due to an obstruction, the car could stop at the side of the road to wait for assistance. On the contrast, cars having to stop at the side of the road would become bothersome for those impatient, but it is intended only for larger ethical decisions or problems where the car has no solution. For example, if a large portion of a road is blocked due to a landslide or another anomaly, then the car would have to decide whether to break rules for an undetermined amount of time, or fail to proceed further. If the self driving car receives instructions from the occupant of the car or even a outside source such as auto assistance, what happens if the provided action to the car results in a disaster, who becomes at fault?

iv.vii Liability of no fault

No fault methodology could succeed in solving how to deal with large ethical dilemmas, but fail regarding smaller ethical decisions. Self-driving cars for the most part will obey the law exactly as programmed, but there could be a case where a self-driving car does not. For example, in construction zones, the car can be ticketed for speeding if it did not account for the change. Who is to blame for the speeding ticket? The owner of the vehicle who has failed to inform the car of construction zone or be on the alert for traffic speed changes, the programmer who did not set up proper ways to identify construction zone changes, or the construction site who did not make the area properly known for being a construction site to self-driving cars. It is also possible that a self-driving car fails to stop at a red light against its will for any number of reasons, such as external factors like bad weather, bad interpretation of light color, or car hijacking, but the car is still ticketed for failing to stop at a red light. With any driver (self-driving or human), cars can have a hard time stopping on ice and end up sliding into the middle of the intersection, does it become the fault of the driver for going out in bad weather or the car for not taking the proper precautions to ensure a safe stop, or the car not having better prepared for such weather? Simply waiving the ticket as it was a mistake is not ethically feasible. Self-driving cars will, in theory, already bring down the rate of tickets being issued as a whole because they are programmed to be law abiding drivers under any circumstance. A lot of police rely on funding from drivers breaking the law as police departments around the United States are faced with higher expenses and smaller budges and have started using traffic tickets as a source of revenue as a result implemented "unofficial" ticket writing quotas. (AOL Autos Staff, 2013) If traffic tickets are currently a way to balance the books for police departments, self-driving cars keeping the roads more ethical will most likely because the cost of the few infractions that do still happen to dramatically increase which will make it harder to convince people to pay for the infraction.

V Further Exploration Ideas

1) ii Assume all users are stupid Establish the 4 laws as a baseline for AI
2) ii Ethics of sentient AI
3) iii.i Expand perfect world ideas: how cars would work in perfect, but more importantly how to handle failures of non perfect world
4) iii.iv Higher standards for AI vs humans
5) iii.vi Correlation vs causation of future prediction and crime rate
6) iii.ix Guilty upon association
7) iv.iii Google, facebook, dataminers with privacy vs safety

VI Conclusion

Self-driving cars can and will bring change to the world in a similar way as other technological advancements like the internet, or space travel. However, with self-driving cars there is potentially greater risk of human harm on a larger scale compared to the technological advances in the past. In order to progress, there needs to be a shift in society's attitude to embrace the technology. This will be non-trivial due to the global scale of adoption, and is not as simple as forcing citizens to accept the technology through laws. Insufficient questions have been asked pertaining to how pedestrians get into vehicle related accidents. Even if these questions are asked before an accident occurs, when the collision results in a fatality, it is impossible to ask the victim how the situation developed.

In vehicle accidents, humans often have the opportunity to identify what societal changes would prevent the accident from repeating, but rarely apply this thinking on a grand scale. Standardized testing to get a driver's license does not include questions such as “if given the option to harm yourself or a pedestrian, which do you choose?” In order for self-driving cars to be universally accepted, questions like these will have to be addressed. Even though humans are expected to make a split-second judgment call in a crisis, self-driving cars will need to be told how to respond ahead of time. Even though self-driving cars will not operate perfectly, they are expected to drive more reliably than the average human, thus reducing the potential for accidents.

References

(Abby Ohlheiser 2017), “The creator of Godwin’s Law explains why some Nazi comparisons don’t break his famous Internet rule” Washington Post, August 14th 2017 https://www.washingtonpost.com/news/the-intersect/wp/2017/08/14/the-creator-of-godwins-law-explains-why-some-nazi-comparisons-dont-break-his-famous-internet-rule/?noredirect=on&utm_term=.ec6dcc45f338

(André Blais 2001), Louis Massicotte, Antoine Yoshinaka, "Deciding Who Has the Right to Vote: A Comparative Analysis of Election Laws," Electoral Studies Vol. 20 (2001), pp. 41–62.
(http://www.elections.ca/content.aspx?section=vot&dir=bkg&document=ec90785&lang=e)

(AOL Autos Staff, 2013) "Police Across USA Charged With Traffic Ticket Quotas To Meet Budgets." Autoblog. Sept. 6, 2013. (April 10, 2015)
http://www.autoblog.com/2013/09/06/police-usa-traffic-ticket-quotas-budgets/
https://people.howstuffworks.com/do-police-write-tickets-to-make-money.htm

(Beall Abigail, 2017) “Driverless cars could let you choose who survives in a crash” *New Scientist magazine 12 October 2017 Journal reference: Artificial intelligence and Law, DOI:10.1007/210506-017-9211-z*

(Craggs Samantha, 2017) “After 60 years, street hockey will be legal in Hamilton again” CBS Interactive Inc. March 1st 2017
http://www.cbc.ca/news/canada/hamilton/street-hockey-is-legal-again-1.4004816

(Excell Jon, 2017) “Driverless cars learn from humans in Greenwich project” *The Engineer October 11th 2017*
https://www.theengineer.co.uk/driverless-cars-greenwich-project/

(Gellert Chantelle, 2017) Sarkar Atrisha, Citulsky Evan, Khan Wasif “How to combine bottom-up learning approaches with top-down rule-based approaches in defining ethical principles for AI systems?” University of Waterloo, 2017

(Gill Kathy, 2017) “What is the Endangered species act?” Thoughtco December 19th 2017
(https://www.thoughtco.com/what-is-the-endangered-species-act-3368002)

(Giuseppe Contissa, 2017) Francesca Lagioia, Giovanni Sartor “The Ethical Knob: ethically-customisable automated vehicles and the law” *Springer Science+Business Media B. V. 2017*

(Government of Canada, 2017) “Canadian Armed Forces jobs, Apply now” Government of Canada, December 19th 2017
(https://www.canada.ca/en/department-national-defence/services/caf-jobs/apply-now.html)

(Government of Canada, 2017) “Driving in Canada” Government of Canada August 24th 2017

(https://www.canada.ca/en/immigration-refugees-citizenship/services/new-immigrants/new-life-canada/driving.html)

(Grossman David, 2018) “Chinese facial recognition system would fine jaywalkers by text”. Popular Mechanics March 28th 2018 (https://www.popularmechanics.com/technology/infrastructure/a19623846/chinese-facial-recognition-system-would-fine-jaywalkers-by-text/)

(Grossman Michael, 2015) “Are Railroads Liable for Hitting Cars or Pedestrians on Their Tracks?” Grossman Law Offices, September 10th, 2015 (https://www.injuryrelief.com/blog/railroad-liability-for-pedestrians-cars/)

(Hurtado, 2016) Hurtado, Manuel. “The Ethics of Super Intelligence” Hurtado M. 2016 doi: 10.4172/2090-4908.1000137

(Kozier, Erb, 2014) “Fundamentals of Canadian nursing, second Canadian edition” Kozier, Erb, Berman, Snyder, Bouchal, Hirst, Yiu, Stamler, Buck. (2014). Toronto: Pearson.

(Kraft Amy, 2016) “Microsoft shuts down AI chatbot after it turned into a Nazi” CBS Interactive Inc. March 25th, 2016 (https://www.cbsnews.com/news/microsoft-shuts-down-ai-chatbot-after-it-turned-into-racist-nazi/)

(Lafrance Adrienne, 2015) “Self-driving cars could save 300,000 lives per decade in America” TheAtlantic September 29th 2015 (https://www.theatlantic.com/technology/archive/2015/09/self-driving-cars-could-save-300000-lives-per-decade-in-america/407956/)

(Latimer Kendall, 2016) “Red light texting still a problem in Canada, says CAA” CBC News December 28th 2016 (http://www.cbc.ca/news/canada/saskatchewan/red-light-texting-still-a-problem-1.3913965)

(LCBO, 2018) “Legal Drinking Age” Liquor Control Board Of (LCBO) 2018 (http://www.lcbo.com/content/lcbo/en/responsibility/responsibility/legal-drinking-age.html#.Wsk2LYjwYdU)

(Levin Sam, Wong Julia, 2018) “Self-driving Uber kills Arizona woman in first fatal crash involving pedestrian” The Guardian March 19th 2018 (https://www.theguardian.com/technology/2018/mar/19/uber-self-driving-car-kills-woman-arizona-tempe)

(Lima Salge, 2017) Alves de Lima Salge, Carloine. Berente, Nicholas. “Computing Ethics is that social bot behaving unethically?” Communications of the ACM Vol. 80, No. 9, September 2017

(Madrigal Alexis, 2017) “Inside Waymo’s secret world for training self-driving cars” The Atlantic August 23rd 2017

(https://www.theatlantic.com/technology/archive/2017/08/inside-waymos-secret-testing-and-simulation-facilities/537648/)

(Markoff John, 2016) “Should your driverless car hit a pedestrian to save your life?” *KitchenerPost June 24th 2016*
(https://www.kitchenerpost.ca/news-story/6739595-should-your-driverless-car-hit-a-pedestrian-to-save-your-life-/)

(Marshall Aarian, 2018) “Uber Video shows the kind of crash self-driving cars are made to avoid” WIRED Transportation. March 21st 2018
(https://www.wired.com/story/uber-self-driving-crash-video-arizona/)

(Metcalfe John, 2017) “The other side of roundabouts: more crashes” City lab March 3rd 2017
(https://www.citylab.com/design/2017/03/the-other-side-of-roundabouts-more-crashes/518484/)

(Neurosci Behav, 2017) “Using virtual reality to assess ethical decisions in road traffic scenarios: applicability of value-of-life-based models and influence of time pressure” *frontiers spotlight July 5th 2017*
https://www.frontiersin.org/articles/10.3389/fnbeh.2017.00122/full

(NIH, 2018) “Alchol Policy” National Institute on Alcohol Abuse and Alcoholism supported by USA Government 2018
(https://www.niaaa.nih.gov/alcohol-health/alcohol-policy)

(Nir Sarah Maslin, 2017) “Woman Who Drove on Tracks at Fault in Fatal Metro-North Crash, Investigators Find” A18 New York edition July 26th 2017.
(https://www.nytimes.com/2017/07/25/nyregion/woman-who-drove-across-tracks-at-fault-in-fatal-metro-north-crash-investigators-find.html)

(Orf Darren, 2014) “Part of the Job: How Engineers Deal with death on the railroad”. Popular Mechanics April 9th 2014
(https://www.popularmechanics.com/technology/infrastructure/a10450/death-on-the-railroad-16675879/)

(Pennachin, 2007) Goertzel Cassio Pennachin, Ben. “Artificial General Intelligence” Springer-Verlag Berlin Heidelberg, 2009

(Royal Canadian Mounted Police, 2016) “Tobacco” Royal Canadian Mounted Police, Goverment of Canada July 7th 2007
(http://www.rcmp-grc.gc.ca/cycp-cpcj/dr-al/tobacco-tabac-eng.htm)

(Sagakiyotaki-ku Kyoto-ken, 2017) “Hiding from Japanese Ghosts” Image: Kiyotaki Tunnel
(https://www.hidingfromjapaneseghosts.com/kiyotakitunnel.html)

(Service Registration, 2018) “America’s preparedness starts here.” Electronic form with the US Government, December 20th 2016

(http://www.militarydraftregistration.com/)

(Shuriye, 2012) Shuriye, Abdi O. "Understanding Engineering Ethics." IIUM Engineering Journal 12.5 (2012).

(Thomson, 2001) Thomson, Alan J., and Daniel L. Schmoldt. "Ethics in computer software design and development." Computers and Electronics in Agriculture 30.1 (2001): 85-102.

(Tobacco21, 2018) “List of all Tobacco 21 Cities” Tobacco21 state by state 2018 (https://tobacco21.org/state-by-state/)

(USA Government, 2018) “Voter registration age requirements by state” USA Government April 3rd 2018.
(https://www.usa.gov/voter-registration-age-requirements)

(Weiss, 1942) Weiss, Paul. "Morality and ethics." The Journal of Philosophy 39.14 (1942): 381-385.