Douglas Gauck

IFSM 304 7985 Ethics in Information Technology (2158)

December 1, 2015

Paper C2, Week Six

Individual Research

Automation: The Ethical Issues of Self-Driving Cars

contents:

Synopsis

The ethical implications of the deployment of computer-assisted or computer-controlled automobiles is an important issue because of the nearly omnipresent use of the automobile. Even in urban areas with highly developed mass transit systems, interaction with cars is commonplace. Who or what is in control of those vehicles, and the strictures under which they will be operated will affect almost every member of society. Drivers of traditional cars, spectator/drivers in automatic cars, passengers, pedestrians, property owners – each have a stake in the rules that will govern their development and usage.

The car itself is a piece of technology, and society has evolved laws and ethics surrounding its current state of development and usage. State inspections, registration, insurance and emissions requirements, as well as driver training, testing and licensing have all developed in reaction to cope with the power and potential danger of the automobile. Laws designed to insure the roadworthiness of the vehicle and qualifications of its operator are intended not for the benefit of the owner/operator who must meet with them, but to try to protect the safety of others in the environment in which the vehicle will be operated. Together these constitute a social contract; one has yet to be written for this technology.

Enabling autonomous control of a vehicle transfers judgment from the human driver to a set of preprogrammed algorithms — for better or worse. The computed results of the algorithms – and the resultant behavior of the vehicle – will be predicated on ethical decisions assigned during programming. Which standards are used, whether the parameters are fixed or adjustable, and whether they align with society's or the driver's own set of values are just a few of the questions for this emerging technology.

Question 1: Does a self-driving car have a driver?

For a human, securing the right to drive a car requires passing a driver's license test comprised of both quantitative and qualitative exams. This involves learning the meaning of road signs, traffic laws, stopping distances, etc. as well as showing competency in using the controls and ability to safely and effectively control the vehicle. After demonstrating a grasp of the knowledge needed to drive, the final hurdle is having a trained evaluator assess a representative sample of the candidate's actual operation of the vehicle. This utilitarian approach safeguards the general welfare.

Most people are able to pass the tests and become competent drivers. Becoming a good driver requires awareness of the surroundings and situation, good reflexes and good judgment. Human judgment is based on training, past experience, morals, mood, attentiveness, ego, and any number of other mentalities. Machine decision-making is based on programming. Computers can be programmed to accept input from sensors, detect road signs and lane markers, compute stopping distances based on mass and velocity, sense other cars and objects and measure their distance, and react to all this as fast or faster than a human, resulting in a self-driving car that could pass a driver's license exam.

Self-driving cars can therefore be said to have a driver: the computer that is accepting various inputs, processing data and outputting commands to control the vehicle's actions. However, the question of judgment remains.

Question 2: What are the risks?

The obvious risks are the consequences in case of malfunction. With increased complexity, the odds of a malfunction increase. Legally required periodic inspections of the

vehicle are designed to minimize mechanical failures, but there is no process in place to assess the suitability of the software controlling a self-driving car.* Both are engineered by humans, but software is a less tangible, more recent invention and subject to bugs, crashes, and unforeseen incompatibilities. A complete system would be the object of attacks from hackers looking for a challenge or more nefarious groups trying to hijack a car for more sinister aims.

Even a totally robust self-driving car could, like any tool, be misused by the human wielding it. The tremendous number of cars on the road has had a major impact on the environment, and if they are able to drive themselves that number is likely to increase. The effect on society of accepting self-driving cars — of yielding yet another aspect of our lives to automation and computerization — is an interesting sociological topic, especially given the symbolic nature of the car as a means of liberation and freedom.

The fundamental risk is that of giving up control. When a passenger on a bus, train, or airplane, control of the vehicle is given over to a complete stranger, but that stranger is still a human being who has been qualified to drive/pilot the vehicle and whom it is assumed shares the human instinct for self-preservation. This engenders an implicit level of trust. A self-driving car would operate inscrutably, its 'instincts' programmed in bits flowing through circuits. The finished product is the culmination of the efforts of hundreds or thousands of people and a lengthy complicated process; a much greater amount of trust is being demanded in a much more abstract entity. (Lin, 2013)

* While there is no shortage of litigation in cases of manufacturer defects, society seems to rank mechanical failures differently from technical ones: the former are viewed as part of natural entropy, while more culpability is sought for computer glitches. Revenge on the nerds?

Question 3: Should the technology be used?

This is the main ethical question — whether to turn over control of an automobile to a computer. While the decision of whether to trust and use a self-driving car will to some degree be a personal one, the commercial availability of the vehicles will first depend upon their acceptance by the general public. The autos must obviously adhere to traffic laws, but will laws be drafted to mandate their decision-making process? Trust is gained by knowing and agreeing with the priorities that govern behavior, human or otherwise. As Hobson (2015) asks, "if these rules are not enforced by law, who do we trust to create the systems that make these decisions? Are you okay with letting Google make these life and death decisions? What this really means is before autonomous cars become commercial, public opinion is going to have to make a big decision on what's really 'OK' for autonomous cars to do (or not to do)."

Under ideal circumstances, a self-driving car does not encounter situations that bring up ethical dilemmas. All collisions are avoided, and no harm comes to anyone or anything. In these cases the use of a self-driving car is ethical under any framework. The egoist can use transit time for other purposes, the altruist gives up control of the vehicle for greater efficiency, the Utilitarian values the overall benefits of the system, and the Deontologist chooses the safer mode of transport.

It is when considering the less-than-optimal circumstances, when the auto must 'choose' a course of action from amongst undesirable alternatives that ethical questions arise. The Utilitarian approach of minimizing harm is the most widely supported until the prospect of the occupants being of lower priority is considered. "People are in favor of cars

that sacrifice the occupant to save other lives—as long they don't have to drive one themselves." ("Why Self-Driving Cars", 2015)

Protecting the occupants at all costs is not acceptable for a computer or a human driver. Contrasting the Utilitarian view, Goldhill (2015) relates:

> the correct moral action doesn't just evaluate the consequences of the action, but also considers who is morally responsible. Helen Frowe, a professor of practical philosophy at Stockholm University … says self-driving car manufactures should program vehicles to protect innocent bystanders, as those in the car have more responsibility for any danger. "We have pretty stringent obligations not to kill people.... If you decided to get into a self-driving car, then that's imposing the risk."

This approach is consistent with existing expectations for human-controlled vehicles, and introduces the factors of maturity, experience and judgment into the operation of a vehicle. The workings of situational awareness and morality are not fully understood in humans, and are a long way from being programmable into a computer. Yet it is precisely those qualities that are being asked of a self-driving car, and Lin (2013) illustrates the difficulty facing manufacturers:

> it would be an unreasonable act of faith to think that programming issues will sort themselves out without a deliberate discussion about ethics, such as which choices are better or worse than others. Is it better to save an adult or child? What about saving two (or three or ten) adults versus one child?… Again, ethics by numbers

> alone seems naïve and incomplete; rights, duties, conflicting values, and other factors often come into play…. Programmers still will need to instruct an automated car on how to act for the entire range of foreseeable scenarios, as well as lay down guiding principles for unforeseen scenarios…. And it matters to the issue of responsibility and ethics whether an act was premeditated (as in the case of programming a robot car) or reflexively without any deliberation (as may be the case with human drivers in sudden crashes).

The automotive industry has already had to brave the ethical realm with existing automation / assistive technologies such as anti-lock brakes and airbags. These are not under direct control of the driver, and can in very rare cases contribute to injury or even death. Because the safety benefits vastly outweigh the instances of harm, the adoption of those technologies has met little resistance (Knight, 2015). Self-driving car technology, once mature and proven, could be viewed and accepted in the same way. They may not become perfectly safe, but the threshold for adoption might be more properly set as when they have become safer than the average driver. “Every year, 1.2 million people die in car accidents, so … moving forward too slowly with self-driving car technology is an ethical problem on its own.” (D’Onfro, 2015)

References

D'Onfro, J. (2015, October 26). The huge, unexpected ethical question that self-driving cars will have to tackle. Retrieved November 26, 2015, from http://www.businessinsider.com/the-ethical-questions-facing-self-driving-cars-2015-10

Goldhill, O. (2015, November 1). Should driverless cars kill their own passengers to save a pedestrian? Retrieved November 26, 2015, from http://qz.com/536738/should-driverless-cars-kill-their-own-passengers-to-save-a-pedestrian/

Hobson, J. (2015, October 29). The Ethics Of Self-Driving Cars Making Deadly Decisions. Retrieved November 26, 2015, from http://hackaday.com/2015/10/29/the-ethics-of-self-driving-cars-making-deadly-decisions/

Knight, W. (2015, July 29). How to Help Self-Driving Cars Make Ethical Decisions. Retrieved November 26, 2015, from http://www.technologyreview.com/news/539731/how-to-help-self-driving-cars-make-ethical-decisions/

Lin, P. (2013, October 8). The Ethics of Autonomous Cars. Retrieved November 26, 2015, from http://www.theatlantic.com/technology/archive/2013/10/the-ethics-of-autonomous-cars/280360/

Why Self-Driving Cars Must Be Programmed to Kill. (2015, October 22). Retrieved November 26, 2015, from http://www.technologyreview.com/view/542626/why-self-driving-cars-must-be-programmed-to-kill/