Computing, Legal, and Moral Ambiguities of Self-Driving Automobiles

Graham P. Wright

CS 3790

Dr. Rose Arriaga

Computing, Legal, and Moral Ambiguities of Self-Driving Automobiles

Artificial intelligence, formally founded in the 1950s, has provided applications to problems spanning a wide range of disciplines such as stock trading, video game playing and data mining, often greatly outperforming humans in terms of reaction time, computation, and knowledge. To the daily commuter, however, artificial intelligence has shown the most promising potential to revolutionize the way we interact with our automobiles. Emerging technologies such as those found in the Mercedes Intelligent Drive and the experimental Google Self-Driving Car are paving the way for a world inhabited by artificially intelligent automobiles capable of replacing human drivers. This integration of artificial intelligence into the daily lives of hundreds of millions of people will serve as a key precedent for future artificial intelligence. As a precedent, though, there are many concerns to address before the realization of such a world: how to program such a complex decision-making program, who is responsible for harm caused by an autonomous system, and how to justify the behavior of a self-driving car in an ambiguous moral situation.

SELF-DRIVING CAR CHALLENGES

Multiple companies have tackled the problem of implementing an assistive driving agent, and Mercedes Benz’s Intelligent Drive showcases examples of many consumer-ready assistive driving technologies. An agent resides in each technology, mapping percepts from the environment (roads, traffic, weather) to actions on the car (acceleration, braking, or steering) through actuators (brakes, steering mechanisms) (Friedenberg and Silverman, 392-393). Assistance comes in the form of automatic control of high beams in traffic (Adaptive Highbeam Assist Plus), identification of parking spaces with automatic parking steering (Active Parking Assist), automatic maintenance of ideal distance behind the car ahead in traffic and lane centering (Distronic Plus with Steering Assist), intervention of one sided braking in the event of a collision-causing lane change (Active Lane Keeping Assist), detection of driver fatigue (Attention Assist), and much more ("On the way…," 2014). This already impressive array of assistive technologies suggests that there are even more to come. The next step is to completely transfer control of the automobile from the hands of a dangerous human to the sound, intelligent agent.

Challenges in Computing. The jump from an assistive intelligent automobile to an entirely autonomous intelligent automobile is where the capabilities of interacting artificially intelligent agents are truly tested. These autonomous vehicles must know how to operate independently with human drivers on the road, although in the future they may rely on communications or "pings" from other autonomous vehicles. There must also be a safe way to pass control to the human riding in the vehicle in the event of an emergency or unknown situation, perhaps creating a dangerous transitional period. Finally, autonomous vehicles must be able to make rational decisions from both fuzzy and novel percepts; these situations include those involving hazardous conditions (what if the car sensor cannot "see"?), traffic cops, unexpected construction, new signs, and more. In essence, the agent must be able to match the human brain's unique ability to incorporate novel information into its knowledge base and respond in the most optimal way, even overwriting previous information if needed.

Agents that can learn in this way are referred to as learning agents, agents that are capable of altering their mapping of percepts to actions (Friedenberg and Silverman, 393). One example of this behavior could be that drivers on a specific San Francisco highway travel 5mph faster than other highways, so the self-driving car realizes this and uses this data to maintain a faster speed on that highway. The car has altered its mapping of percept (traffic flow) to action (acceleration), analogous to the human brain's ability of plasticity. A puzzling sub problem is the method by which an autonomous car can emulate the wide range of cognitive abilities of the human brain.

Legal Challenges. The next obstacle in the realization of widespread automotive autonomy concerns the legalities of an autonomous agent. In May 2013, the National Highway Traffic Safety Administration (NHTSA) introduced policies targeting autonomous vehicles including five levels of vehicle automation: from No-Automation (Level 0) to Full Self-Driving Automation (Level 4) (Aldana, 2013). A car fully equipped with Mercedes Benz Intelligent Drive would fall in Level 2, Combined Function Automation, and the Google self-driving car would fall in Level 3, Limited Self-Driving Automation. Already, states such as California, Florida, Texas, and Nevada have passed legislation that permits the testing of self-driving vehicles, but this is only the beginning of a longer assessment required to understand the safety measures required in a world filled with autonomous automobiles (Kirkpatrick, 18). Other considerations include the revamping of the driving test, the training required of existing drivers, the safety regulations of self-driving cars, and the altering of traffic laws to allow for completely autonomous vehicles.

Unlike a master chess-playing agent, autonomous cars will bear the responsibility of transporting its passengers to a destination with the least harm. The world has seen artificially intelligent agents for years, but the self-driving car is a legal precedent for the determination of responsibility in an accident involving an autonomous agent. This introduces a debate on a fundamental cognitive science topic: who is responsible for an artificially intelligent agent's mistakes? At what point should a human have overridden these potential mistakes? Many ambiguities arise from these thought experiments, requiring very precise enumerations in law. Automobile insurance must also be reworked as driving responsibilities shift from man to machine. These are relatively inconsequential hurdles considering the number of deaths that may be prevented. According to the National Highway Traffic Safety Administration, in 2012, 33,561 deaths resulted from automobile related accidents, averaging 10.69 deaths per 100,000 people ("National Statistics," 2012). This number could see a sharp decline in the future with the increasing safety of vehicles and introduction of assistive driving technologies.

Moral Ambiguities. Autonomous automobiles must be able to act rationally in situations that are ambiguous due to less precise sensor information, but how does a predominantly objective agent respond in a morally ambiguous situation? A classic example of a morally ambiguous problem is the trolley problem. In this thought experiment, a man has the choice of either letting a runaway trolley run over and kill five people or pulling a lever that redirects the trolley to run over and kill just one person. Similarly, an autonomous vehicle might encounter a situation where there is an imminent collision, avoidable only by swerving into another car or pedestrian. This scenario would require an extensive risk analysis algorithm in order to lead to the least harm caused.

Another scenario exposes possible flaws of unintelligent sensing devices. While a human may be able to understand that a floating trash bag poses no threat to his two ton machine, an autonomous car might identify this object as a lethal hazard and swerve to avoid it, perhaps hitting a pedestrian. This type of decisioning system requires a formally accepted code of morals for computing so that a decision made by an autonomous vehicle has an objective justification: "In [some] cases, the robot or autonomous machine may be functioning as intended by the designer without any clear errors or failures, and if unintended results still occur, the issues of responsibility and liability must be sorted out" (Kirkpatrick, 17). Unlike a human operating a vehicle, an autonomous agent relies on the code that it was shipped with to make all of its decisions; there is no excuse of "human error." This requires substantial defense for every decision that the autonomous agent performs in order to minimize legal risks for the engineers behind the technologies.

Trusting the Agent. After conquering the technical, legal, and moral challenges, the final step is convincing the public to adopt the autonomous car. Considerations include cyber security concerns, reluctance to give up driving control, distrust of software, simplicity and flexibility of user interface, and price. Consumers want to be convinced that an autonomous car will be safer than their own driving, just as easy to use as a normal car, and ultimately worth the investment. Safety concerns are very valid, especially in a system as complex as an autonomous automobile. In a world filled with self-driving cars, almost every situation imaginable will test the robustness of the software engine. History has shown that one miscalculation can lead to disaster; NASA suspected that the loss of their Mars Polar Lander was due to a software error that misidentified a vibration as a landing (Kopec and Tamang, 180-184). A safe and reliable self-driving system will require extremely extensive testing in every driving situation, including all rare edge cases. The problem lies not in making a car drive on its own. As of April 2014, Google's self-driving car has logged over 700,000 miles under computer control without an accident (Urmson, 2014). The problem lies in dealing with every possible driving scenario that can occur.

CONCLUSION

Although fully autonomous vehicle technologies have made impressive leaps in the last few decades, there are still a wide range of regulatory, legal, and technical hurdles that must be addressed before a fully autonomous vehicle can be released to consumers. This release is inevitable in the future; a successful solution to this array of self-driving challenges will provide drastic benefits to the public. Less traffic, faster travel times, less traffic police, less need for car insurance, relaxed laws for who can “drive” in a car, higher productivity during traveling, better fuel efficiency, and increased traffic safety are all very realistic advantages of a world filled with autonomous vehicles. Execution is essential in realizing these advantages; all proper precautions, moral, legal, and computing, must be in place to deem the autonomous car a success. The release and popularization of the self-driving car will act as a framework for the future incorporations of even more autonomous technologies into everyday life.

Works Cited

Kopec, Danny and Suzanne Tamang. “Failures in complex systems: case studies, causes, and possible remedies.” ACM SIGCSE Bulletin 39.2 June 2007: 180-184. ACM Digital Library. Web. 30 Sept. 2014.

Kirkpatrick, Keith. “Legal issues with robots.” Communications of the ACM 56.11 November 2013: 17-19 ACM Digital Library. Web. 30 Sept. 2014.

“National Statistics.” Fatality Analysis Reporting System Encyclopedia. National Highway Traffic Safety Administration, n.d. Web. 30 Sept. 2014.

Aldana, Karen. “U.S. Department of Transportation Releases Policy on Automated Vehicle Development” Press Release. National Highway Traffic Safety Administration, 30 May 2013. Web. 30 Sept. 2014.

Friedenberg, Jay and Gordon Silverman. Cognitive Science An Introduction to the Study of Mind. Thousand Oaks: SAGE Publications, Inc., 2012. 389-424. Print.

"On the way to accident-free driving. Mercedes-Benz Intelligent Drive." *Mercedes Benz Intelligent Drive*. Mercedes Benz, n.d. Web. 5 Oct. 2014.

Urmson, Chris. “The latest chapter for the self-driving car: mastering city street driving.” Web blog post. *Official Google Blog*. Google Inc., April 28, 2014. Web. October 15, 2014.