The National Highway Traffic Safety Administration (NHTSA) recently released federal guidelines for manufacturers on the testing and deployment self-driving cars (1). Shortly after, the California Department of Motor Vehicles (DMV) released a revised draft of regulations that essentially allowed manufacturers to test cars with no human in the car on public roads, as long as the manufacturers abided by the federal government’s 15-step assessment guidelines.1

These two developments—along with similar announcements by cities including Boston, Massachusetts; Pittsburgh, Pennsylvania; and Austin, Texas—make clear that driverless cars are a near-term reality. The general public soon will be driving in an environment that includes driverless cars under testing. Theoretically, when these test vehicles generate enough miles to convince state DMVs and NHTSA that driverless cars are safe, then everyday consumers will be able to buy their own driverless cars or to summon a rideshare driverless car.

If driverless cars work as advertised, people with impaired vision or mobility will have new forms of transportation that could change their lives; moreover, the approximately 38,000 annual fatalities caused by driver error could be reduced significantly. These are laudable goals.

Nevertheless, these claims of benefits should be interpreted in the context of the significant amount of money being poured into this industry. Global automotive research and development expenditures are estimated at $94.2 billion for 2016—more than three times the 2016 global expenditures on aerospace and defense (2).

Ready for Deployment?
Industry generally and optimistically predicts that driverless cars will be ready for widespread deployment sometime between 2018 and 2025.2 Two recent fatalities of drivers of Tesla vehicles, one in China and one in the United States, however, have highlighted the complexities and frailties of semiautonomous systems. Both drivers were using a feature known as Autopilot. These fatalities, as well as other related accidents, call into question the capabilities of the systems and the lack of a process for testing and certification.

1 https://www.dmv.ca.gov/portal/wcm/connect/211897ae-c58a-4f28-a2b7-03cbe213e51d/avexpress/terms_93016.pdf?MOD=AJPERES.

In the Tesla fatality in China, a driver in Autopilot mode slammed into the back of a street cleaner on a highway. Five months later, a Tesla driver, also in Autopilot mode, was killed in Florida when the car’s technology failed to detect a tractor trailer turning ahead, and the car, traveling at 74 mph, hit the truck broadside.3

Tesla maintains that both drivers were at fault for not paying attention to Autopilot, technically a “driver assist” technology in a car that is not intended to be fully autonomous. Apparently the drivers did not understand this nuance—perhaps they were misled by the name Autopilot. Similar examples of “mode confusion” are well known in aviation and will only increase as automation becomes more prevalent in cars.

These two fatalities highlight several issues raised before the Senate Commerce Committee in March 2016—namely that the entry of driverless cars into the market may reveal many unknowns, but in the meanwhile, manufacturers are failing to address many problems that are known (3). For example, Tesla knew about the inability of Autopilot to detect static objects on highways, and the owner’s manual warned drivers that the car may not brake for stationary vehicles, especially when the car is driving faster than 50 mph.4 A significant flaw in the car’s perception system—that is, how the car “sees” the world—and the lack of transparency to the drivers led to these two fatalities, and these are problems not easily solved.

Sensors and Postprocessing

The sensors that help driverless cars see can include some combination of radar, lidar or light detection and ranging, computer vision, and ultrasound devices. No single technology can provide complete coverage; because some combination of these sensors must be used, a complex data fusion is required. Moreover, each of these sensors has known limitations, as illustrated by the Tesla fatalities; notably, inclement weather—fog, rain, and snow—has presented problems.

The postprocessing of the data gathered by the sensors requires significant estimating and pattern matching, often referred to as machine learning. As a result, when an expected driving scene does not match the scene observed by the sensors—which may themselves be flawed—an autonomous car may

---


not be able to reason accurately about the world around it and determine the correct next actions in the required time.

These sensor and postprocessing difficulties are widely known in automotive robotics and in all robotics industries that rely on these technologies—including unmanned aerial vehicles or drones, manufacturing robots, and medical robotics. Significant academic and industry research efforts are under way to improve these technologies and processes; however, substantially fewer efforts are developing test strategies to ensure that these stochastic systems work not only in expected driving conditions but also in the boundary conditions in which catastrophic failures occur.

Because driverless cars strongly rely on pattern recognition and probabilistic reasoning, the test strategies that were used for deterministic systems do not work. Because these stochastic systems have embedded complexities, the cars cannot compute a solution to a four-way intersection the same way each time, for example. For driverless cars and for many unmanned systems, an industrywide consensus is lacking on how to test such probabilistic systems to guarantee some level of safety.

Assessing Safety

Driverless car companies have generated a “miles driven” metric to assure safety. RAND Corporation has stated that driverless cars must drive 275 million miles without a fatality to prove that these cars are as safe as human drivers, at a 95 percent confidence level (4). Tesla logged 130 million miles before the U.S. fatality—the most of any company—falling significantly short of the RAND Corporation metric.

Miles driven is not an acceptable solution for demonstrating that a technology is safe for public roads, especially when the numbers are generated in sunny climates with white lines clearly visible on well-maintained highways. Tests should challenge the stochastic reasoning of these autonomous systems, as well as the responses of the cars in the corner cases—the worst possible scenarios the cars could encounter—including snow, ice, fog, environments dense with pedestrian and bicyclist traffic, and unexpected maneuvers from other cars.

NHTSA’s guidelines lay out a 15-point plan for states to follow in assessing whether driverless car technology is ready for use on public roads (1). This assessment plan only addresses high-level areas of concern, such as privacy, system safety, and object and event detection and response. The NHTSA plan does not offer guidance or assistance on assessing each of the 15 areas, leaving each state to interpret and perform its own evaluations, which no doubt will vary widely.

The evaluation of driverless cars is extremely difficult and requires engineers who are experts in both the hardware and the software, as well as in artificial intelligence. In its guidelines, NHTSA admits its lack of staff qualified to make the assessments and suggests that it may develop a network of experts for help in understanding the issues.

No other plans are under way, however, to centralize or to disseminate expert knowledge at either the federal or state level. As a result, state governments will be expected to acquire the expertise to assess the validity and comprehensiveness of driverless car test plans in a short time, despite the lack of commonly accepted standards or a consensus on how to conduct such testing. Because the testing and evaluation of autonomous systems is a nascent field with little foundational research, either theoretical or empirical, the expectation that state governments can do what researchers have not yet demonstrated is a tall order.

Informed Consent

Without a principled approach to the testing of autonomous systems, the implications for the general public are not clear. California soon will allow tests of driverless cars on public roads, with a remote operator monitoring the system. This raises the important issue of informed consent for the public. Although NHTSA’s 15-point plan sanctions such tests, the guidelines do
not address the applicability of the federal regulation mandating that all humans involved in an experiment should explicitly give their consent.

Should drivers be given the option to share the road with one or more driverless vehicles undergoing testing, especially without safety monitors? These cars have no established minimum safety standards, and the state evaluators who would determine road worthiness and public safety are not likely to have the appropriate background to make the judgment. At a minimum, discussion is warranted about clearly marking the driverless cars that are undergoing testing, so that drivers who are sharing the road have some understanding of the test environment for which they did not volunteer.

**Areas for Research**

What do these issues mean to the research community? The promise and potential benefits of driverless cars will be transformative, but further research and development is needed as the rush to deploy driverless car technology has outpaced the technical underpinnings. Significantly more research is needed in a range of areas, including sensor development, artificial intelligence and machine learning, the testing and evaluation of autonomous systems, and the legal, ethical, and public policy implications of driverless cars.

More interdisciplinary work is needed across these fields to communicate the capabilities and limitations of these probabilistic systems. For example, more work is needed in explainable artificial intelligence, to understand the best ways of communicating the outcomes of machine learning algorithms to researchers and policy makers. The Tesla fatalities highlight the gap between the engineers who design complex systems and human users who do not understand the systems.

Human–robot interaction is another major area for research to ensure that reciprocal intent is communicated effectively between all entities within the sociotechnical systems of driverless cars—including the cars, the human operators, pedestrians, bicyclists, and others. Broader sociotechnical questions include the effects on public transportation, the fuel types and requirements of the vehicles, and the effects of projected demand on air quality and congestion.

**Educating a Workforce**

Universities and colleges need to increase the numbers of students entering these fields—the demand exceeds the supply of electrical, mechanical, and computer engineering students and of software developers, who are the core of the driverless car community. The growing sociotechnical issues call for the development of programs that address the interdisciplinary aspects of driverless cars—government and industry badly need graduates with that knowledge and expertise. Universities must adapt to these increasing demands, as must government agencies and the foundations that provide scholarships and incentives for relevant new programs.

Educating a multidisciplinary robotics workforce is critical both in the United States and worldwide; driverless cars represent only one rapidly growing robotics industry. Commercial drones, manufacturing robotics, medicine, and other industries attempting to introduce more autonomous operations are competing for the same people, and the chokepoint currently resides in higher education.

**References**