# AUTONOMOUS SYSTEMS DESIGN, TESTING, AND DEPLOYMENT: LESSONS LEARNED FROM THE DEPLOYMENT OF AN AUTONOMOUS SHUTTLE BUS

*Lance Sherry, John Shortle, George Donohue, Brett Berlin, Jonathan West*
*Center for Air Transportation Systems Research at George Mason University, Fairfax, Virginia*

## Abstract

Advances in technology have enabled the deployment of unprecedented levels of automation that verge on completely autonomous systems such as unmanned passenger and cargo vehicles, and air traffic control supported by integrated communications, navigation and surveillance (ICNS) systems.

One application of the new technologies is in autonomous shuttle buses. This paper describes an analysis of a collision between an autonomous shuttle bus and delivery tractor-trailer on an urban street in Las Vegas. The analysis provides lessons learned for the design, testing, and fielding of future autonomous systems. First, the analysis demonstrates the difficulty in designing for all the "corner-cases" for safe fielding of an autonomous system. Second, the analysis shows the difficulty in demonstrating safety compliance to a target level of safety for systems developed using machine learning that cannot be tested using traditional testing methods (e.g. code-inspection or forms of input-output testing. Third, the analysis identifies the need for the explicit, intentional design, not an afterthought, of the task of the "safety driver." Solutions to these three issues are discussed.

## Introduction

Advances in technology have enabled the deployment of unprecedented levels of automation approaching near-autonomous systems such as unmanned passenger and cargo vehicles, and air traffic control.

One example of an application of autonomous systems is in urban, fixed route, shuttle buses.

Traditional guidance and control systems (G&CS), used in many types of transport vehicles, command a vehicle to follow a pre-defined path using state information derived from sensors such as Global Positioning Systems (GPS), Inertial Navigation Systems (INS), and/or radio navigation aides [1], [2]. These designs are sufficient for vehicles that operate in a near-sterile environment in which progress on the pre-defined path can be managed by basic sensors on the vehicle.

To operate in a non-sterile environment, such as an urban street, the vehicle needs information about the environment such as obstacles, traffic, and flow instructions (e.g. traffic lights). In one implementation, a complex, ubiquitous surveillance and communication infrastructure is deployed. This external infrastructure provides real-time communication to the vehicle on the current and future state of the environment.

In an alternate, "autonomous system" implementation, the vehicle is not dependent on an external, infrastructure derived communication, navigation and surveillance infrastructure. In this implementation, the vehicle can derive the necessary information through its own suite of sensors. In this way, the vehicle can be "dropped in" to the existing environment and is capable of operating autonomously.

Recent advances in new reliable, inexpensive sensors, such as cameras, LIDAR, and radar, along with advances in Machine Learning, make it feasible to fuse multiple sources of sensor data and develop algorithms to commands complex vehicle guidance and control behavior. With this capability, vehicle guidance and control systems (G&CS) can now detect obstacles, traffic, and traffic flow instructions on their own, and make decisions on how to respond to complex, non-sterile environments.

Traditional vehicle G&CS for operations in sterile environments exhibit finite and relatively low complexity. As a consequence, the G&CS can be coded using a combination of rule-based and continuous closed-loop functions [3]. The intended

behavior can be specified a-priori and tested completely (i.e. every combination of inputs has an output), and comprehensively (i.e. every output is determined by a valid combination of inputs). Established methods for verification testing, validation testing, and demonstrating safety compliance are applied and used for regulatory approval. These systems are certified to 10-5, 10-7, or 10-9 target levels of safety [4].

Due to the complexities in recognizing non-uniform objects and situations in the non-sterile environment, and in responding to the exponentially large number of combinations of events, vehicle G&CS for non-sterile environments are "coded" using Machine Learning supervised-training techniques (instead of the traditional rule-based software algorithms). In supervised-training, the Machine Learning "rides along" with a human operator recording their response to every situation experienced by the vehicle. After experiencing the response to a specific situation enough times, the Machine Learning algorithm can "learn" the correct response and distinguish the correct situation/response from other similar situations/responses.

To deploy safe and secure autonomous G&CS for a shuttle bus, an appropriate response to all the possible emerging scenarios that can occur on an urban street must be encoded into the vehicle G&CS. In a dynamic, complex, and ever-changing domain such as urban streets, there can be hundreds of thousands of static and emerging situations that can occur. Before being deployed for revenue service operations, the vehicle G&CS must be tested to demonstrate that all of the possible emerging scenarios that can occur in the field do not result in a hazardous outcome [5], [6].

Modern engineering best practices attempt to design vehicle G&CS to address all the possible scenarios that can occur in the real-world. Due to the complexity of the domain, and the combinatorics, it is possible to deliver a G&CS that does **not** cover unusual situations, known as "corner-cases." One way to uncover missing corner-cases, and to demonstrate safety, is to accrue tens of thousands of hours of operation with an "attendant" or "safety driver," monitoring the vehicle G&CS with the responsibility to intervene should a situation emerge with a potentially hazardous outcome [7].

Monitoring and intervention for hazardous rare events is a complex human-machine interaction process that human operators are not well suited to perform without careful design of procedures and the associated vehicle G&CS user-interface [8] [9].

This paper describes an analysis of a collision between an autonomous shuttle bus and delivery tractor-trailer on an urban street in Las Vegas to provide lessons learned for the design and testing of future autonomous systems [10].

The analysis identified three main lessons learned for the design, testing and fielding of autonomous systems.

First, the design of the guidance and control system must anticipate and handle all the possible real-world situations that can occur. Even on a simple shuttle bus loop route, an unusual situation presented itself 7 minutes after the start of deployment, that was nor properly coded in the autonomous control systems. This emphasizes the difficulty in designing for all the "corner-cases" for safe fielding of an autonomous system and emphasizes the need for alternate methods for enhancing the scenarios that the vehicle is exposed to in the supervised learning phase of the G&CS development.

Second, the use of Machine Learning in the design of the G&CS results in "black-box" automaton that prohibits the utilization of traditional testing methods, approved by regulators, to demonstrate compliance with safety requirements. As a consequence, the Machine Learning G&CS cannot undergo a code-inspection or any form of input-output testing as the input-output relationship is "hidden" in the machine learning code algorithm.

An alternative means of compliance is via performance/risk-based testing in which the system performance is recorded during demonstration testing and used as evidence for safety assurance. This approach is problematic. The number of miles that is required to demonstrate safety may only be completed after several years, possibly after the technology is already obsolete. For example, if the US driving fatality rate is 1E-8 per mile, then autonomous vehicles need to be 100 times safer (i.e.

1E-10 per mile). The statistical "rule of 3" says that if N data points are observed with zero fatalities, then the 95% upper bound on the fatality estimate is 3/N. So that would require 3E-10 miles with zero fatalities to demonstrate 1E-10 per mile target level of safety. If a manufacturer has 1E-7 miles driven to date, then they need to repeat the testing completed to date 3,000 times (with no fatalities).

Third, to mitigate the potential for missing situations in the autonomous G&CS that lead to hazards during the performance-based testing (described above), the regulators and operator must insert a human operator with responsibility to intervene in the event of a hazardous situation. The human operator's role, however, must be an intentional design, not an afterthought. The role and responsibilities of the human must be explicitly designed and supported by appropriate user-interfaces and the limitations of human reliability in monitoring for rare-events must be considered.

The implications of these lessons for the design and deployment of future autonomous shuttle bus systems and aviation systems are discussed.

This paper is organized as follows: Section 2 describes the design and operation of the autonomous shuttle bus, the vehicle guidance and control system, and the attendant. Section 3 describes an accident scenario that occurred following deployment of an autonomous G&CS. Section 4 discusses the issues and lessons learned from the deployment of the autonomous G&CS and the way it supports the safety-driver role. Section 5 concludes with implications of this analysis, concepts to mitigate these issues, and future work.

## System Components

The system had the following components: the shuttle, the autonomous shuttle guidance and control system, and the attendant and their manual control user-interface

### *The Shuttle Bus*

The shuttle bus is designed to transport a total of 15 passengers, 11 seated and 4 standing. It is a two-axle, battery-powered automated test vehicle with a Gross Vehicle Weight Rating (GVWR) of 3,500 pounds (Figure 1). The shuttle had two symmetrical ends, either of which serve as the front

**FIGURE 1: Shuttle bus involved in the collision. Passenger entry/exit through sliding side door. Seating for 11 and standing room for 4. Can operate in either direction forward/backward.**

or the rear. Passengers enter and exit through double sliding doors on one of the long sides of the vehicle.

The shuttle bus is powered by two batteries. An 80-volt traction battery operates the vehicle's electric motor. This battery is located at one end of the shuttle and accessed from an external door. A 12-volt battery serves as a back-up for operating the doors and other miscellaneous non-propulsion functions. The backup battery is located in an enclosed space inside the passenger compartment (at the same end as the 80-volt battery).

Two emergency stop buttons are located on either

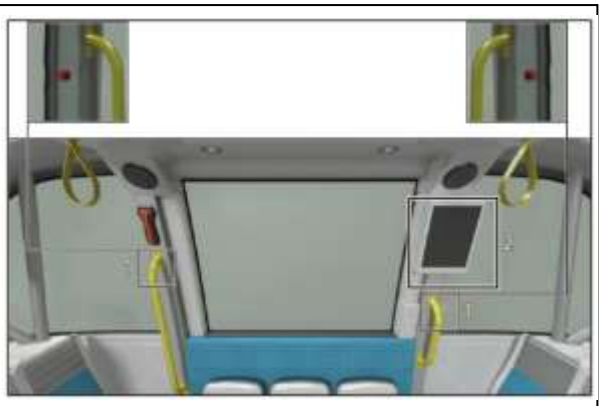

**FIGURE 2: On long-side opposite door: (1) red emergency stop buttons (inset 1), (2) information display screen located (inset 2).**

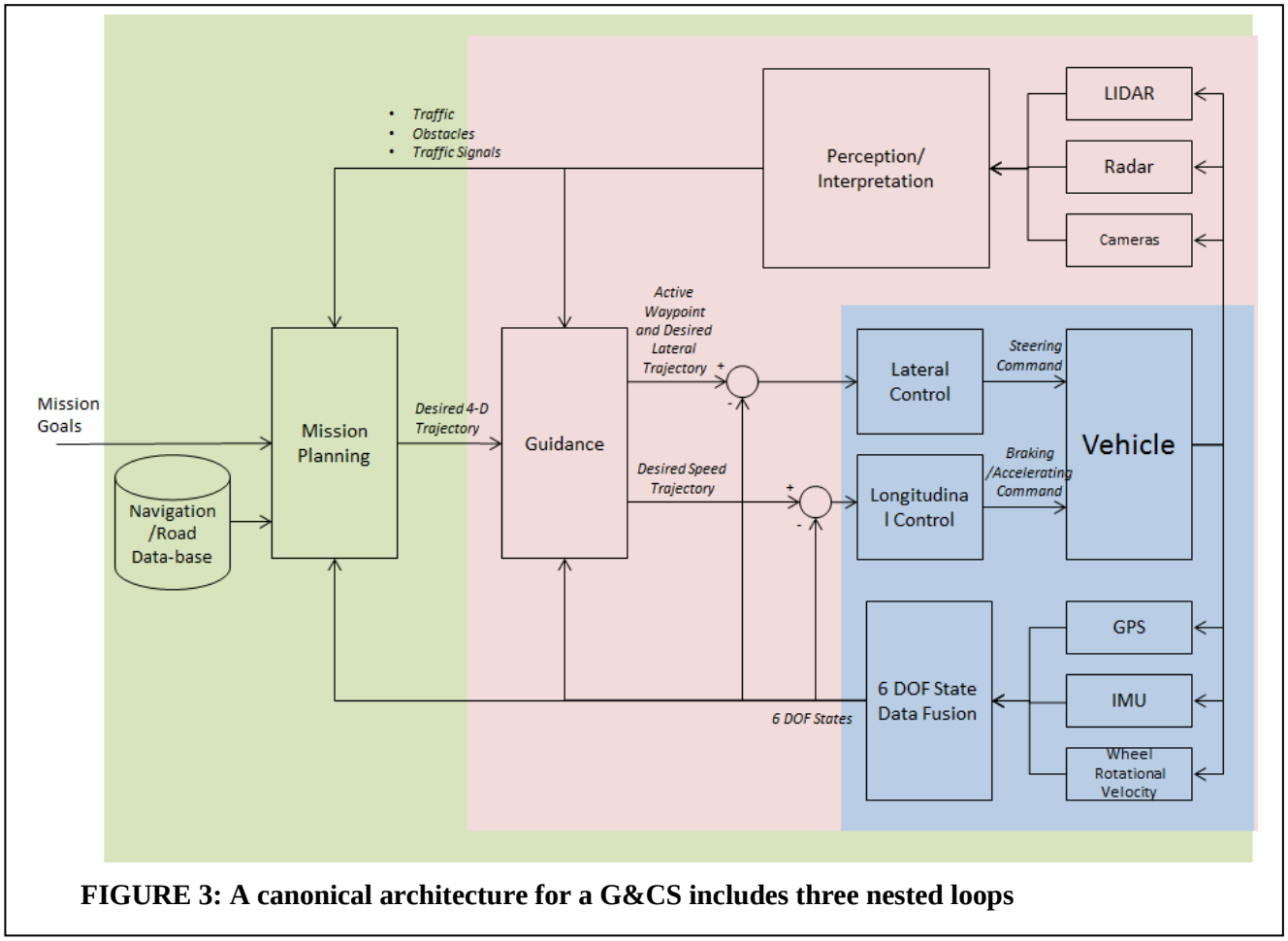


**FIGURE 3: A canonical architecture for a G&CS includes three nested loops**

side of the central window opposite the loading doors (Figure 2). Pushing an emergency stop button turns off the motor, activates three types of brakes, and turns on flashing hazard signals.

A navigation touch screen is also located on the long side opposite the loading doors (Figure 2). The screen displays information such as the battery's charge status and the vehicle route.

The loading doors are equipped with emergency release handles. If the doors cannot not be used to evacuate the shuttle, a hammer is available to break the central window, marked "emergency exit."

A fire extinguisher and first aid kit are stored under the seats opposite the doors.

As the vehicle is designed primarily for autonomous operation, the shuttle does not have a steering wheel, a brake, or an accelerator pedal.

### *Shuttle Bus Autonomous Guidance and Control System (G&CS)*

The Shuttle Guidance and Control System (G&CS) commands the shuttle along a pre-defined path defined by a sequence of legs/waypoints in a navigation data base. The pre-defined path identifies the latitude and longitude for each point on the route and speeds or turn radii on each segment.

To operate in an urban setting, the G&CS must also have knowledge of the obstacles (e.g. pedestrians, work zones, …), traffic, and traffic flow instructions (e.g. such as stop signs, traffic lights) and roadway features (such as grade).

A canonical architecture for a G&CS includes three nested loops (Figure 3). The inner-nested loop is path-tracking closed-loop Control for speed control (I.e. accelerator, brake) and direction (i.e. steering). The function can be accurately designed

and coded using tradition rule-based/continuous functions.

The desired path is determined by the next outer-loop known as Guidance function. This function used GPS, IMU and rotational velocities to determine the vehicle's dynamic state, but supplements this information with non-sterile environment information. The waypoint or segment is adjusted to account for roadway conditions, traffic signals, traffic or other obstacles on the desired path. The Guidance function can change the velocity on the desired path, and/or perform lateral path offsets.

The outer-loop adjusts the mission for contingencies. For fixed-route shuttle bus operations, the Mission Planning function is not invoked.

The Shuttle G&CS has limits on how much it can deviate from the designated route. In the event, the shuttle must deviate from the route, an attendant must disengage the autonomous control system and the maneuver the vehicle using manual controls or modify the mission plan.

### *Sensors*

To keep the shuttle on it's designated path, the shuttle includes a differential Global Positioning System (GPS) to identify the vehicles latitude and longitude, an Inertial Measurement Unit (IMU) measures the shuttle's velocity, acceleration and angular rate to refine its position and verify its location. Also, an Anodometry device measures the speed of the wheels to estimate changes in the vehicle's position.

To keep the shuttle from colliding with objects in it's path and to assist in location identification, the shuttle has eight LiDaR (light detection and ranging) sensors and two stereoscopic cameras. LiDaR measures the distance to other objects using a laser and has a detection range of 40 meters under ideal conditions. Two LiDaRs are positioned on the roof to give a 360-degree view around the vehicle. The primary purpose of the LiDaR is to detect obstacles, whether moving or stationary (cars backing out of parking spaces, motorcycles, bicycles, pedestrians, and so forth) on the roadway or sidewalk. The LiDaR are also used to verify the shuttle's location and path by matching objects and features. The stereoscopic cameras were mounted on the shuttle to monitor the outside environment as well as to analyze signs and traffic signals.

The sensors and their locations are illustrated in Figure 4.

The shuttle also has a dedicated short-range communication system and a long-term evolution antenna to communicate with traffic signals along the route.

Another camera (fish-eye) was mounted on the

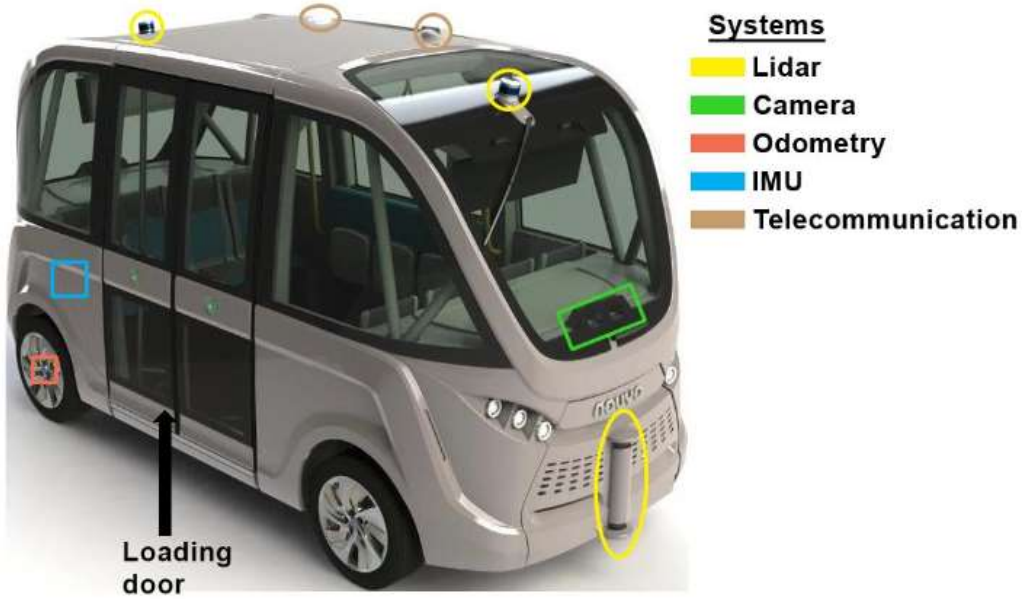


FIGURE 4: Sensors and communication packages on the Shuttle.

ceiling of the shuttle to monitor passengers.

The shuttle's performance was monitored in real time from a control center in Lyon, France. The control center operated 24 hours a day, 7 days a week.

The designated path is created during a map-making activity. The LiDaR and camera systems record environmental features such as roadway markings, curbs, stop lines, traffic signals, signs, road grade and curvature, and certain nontraffic static objects such as buildings. As the shuttle travels along the mapped coordinates of its path, the various systems continuously scan the environment and verify that the detected objects (e.g stop signs) and roadway features (e.g grade) match those on the mapped route at those specific locations.

### *Attendant*

Although the shuttle can operate autonomously, regulatory authorities require an attendant to be on board to supervise and intervene if necessary.

According to shuttle's operator training booklet, the attendant's duties include:

(1) receiving passengers on board

(2) checking that the vehicle functioned properly

(3) reporting errors to the supervision center

(4) maintaining the security of passengers inside the vehicle and of pedestrians outside

(5) reporting damage or injuries

(6) monitoring and intervening as necessary

The attendant is responsible for initiating the shuttle's autonomous operation, can request stops at designated locations, and opens and closes the doors. Only the Guidance and Control System can initiate departure one the system is turned on.

In the event of an unexpected or erroneous action by the shuttle's autonomous system, the attendant can notify Navya by pressing an intercom button on a speaker next to the navigation touch screen. Activating the intercom connects the attendant to the control center in France.

Part of the attendant's regular duties included using the hand-held controller to load the shuttle on and off a tow truck (to take it to and from its route location), to maneuver the shuttle in the yard where it is stored, and to manuever the shuttle into parking spaces.

The attendant is also required to operate the shuttle manually if an obstacle blocked its path. The does not deviate outside its designated path (for example, a stalled vehcile). In these situations, the attendant uses a handheld controller to maneuver the shuttle around the obstacle and then return it to its path. The attendant then re-engage autonomous mode.

### *Manual Controller Used by the Attendant*

A trained driver (attendant) can use manual control to operate the shuttle outside the vehicle's predetermined path (for example, to move it from a storage location to its mapped route or to navigate around stationary objects). This operation is accomplished using an X-Box-style hand-held controller (Figure 5).

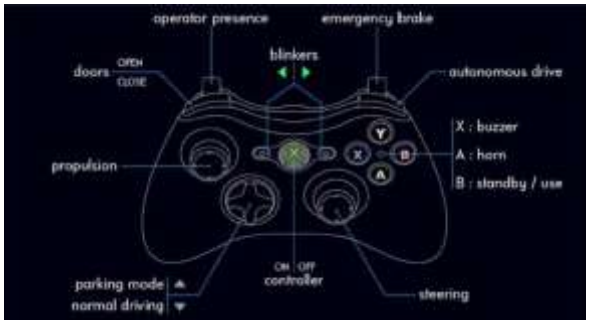


**FIGURE 5: Hand-held controller for manual operation of the shuttle**

Pressing the "operator presence" button on the controller activates manual mode. In addition to steering the shuttle, the controller engages the emergency brake, horn, or buzzer; opens/closes the doors, activates the turn signals (blinkers). Pressing both turn signal buttons activates the hazard warning lights.

Releasing the control button (green X at center of controller) activates the emergency brake. Pressing the "standby" button disables propulsion.

Pressing two buttons on the controller ("operator presence" and "autonomous drive") returns the shuttle to autonomous mode.

Prior to this incident, the controller was stored in an enclosed space at one end of the passenger compartment. In the incident, the attendant did not retrieve the controller during the event. After the crash, a new company policy was established such that Attendants now remove the controller from its storage space at the beginning of a trip and keep it available throughout the trip.

### *The Accident Scenario*

The shuttle was a test vehicle, part of a pilot program in Las Vegas, and was on its first day of passenger-carrying operation (shuttle rides were free) when the collision occurred. The designated path is 0.6 mile circular route with right turns only. The vehicle was limited to a maximum speed of 16 miles per hour.

The Shuttle started the trip at Container Park on Fremont Street heading west (Figure 6). At the beginning of the trip, the attendant boarded with the passengers at Container Park and started the shuttle in autonomous mode. Eight seats were occupied at the time of the collision.

The shuttle then turned right onto South 8th Street, then right again onto East Carson Avenue. It stopped at an information kiosk at South 8th Street and at East Carson Avenue.

The Shuttle then turned South onto South 6th Street. A little over halfway down South 6th Street there was a tractor-trailer delivery truck backing into an alley.

The truck driver was backing into an alley perpendicular to South 6th Street. This maneuver is a standard procedure to back into the alley as the truck was "too long" to enter the alley from Las Vegas Boulevard. Backing into the alley is considered the "only way . . . to be safe pulling out."

When the shuttle bus turned on to South 6th Street, the tractor trailer, facing south, had pulled up past the alley. Before backing up, the driver told investigators that he activated his flashers as he pulled up to the alley. Two cars were behind him. He waved the two vehicles by.

When the driver began backing up, he saw the shuttle turn onto South 6th Street from Carson Avenue. He said that he knew the shuttle was automated having seen it previously doing "test runs" on Fremont Street. He said that he had no concerns about sharing a road with the shuttle and assumed it would come to a stop to allow the truck to enter the alley.

As the driver backed-up he paid particular attention to vehicles parked on the east side of South 6th Street so as not to strike them. The cars were on the left side of the truck.

At this time, the driver looked to the right and noted that the shuttle was halfway down the street. He stated that he assumed the shuttle would stop a "reasonable" distance from the truck.

The driver said that he looked back to the left and saw a pedestrian in the alley. He waited until the pedestrian cleared.

During this time, the attendant told estigators that the shuttle slowed as the shuttle roached the tractor-trailer. The shuttle stopped meters (10.2 feet) from the tractor-trailer.

Just before the shuttle stopped, the attendant, ure if the shuttle would come to a complete stop, ssed one of the emergency stop buttons on the l opposite the loading doors. Recorded data ws that the shuttle's speed was less than 1 mph 249 meters per second, or 0.56 mph) when the ndant pressed the emergency stop button. This on disengaged the autonomous G&CS. The shuttle was now under manual control.

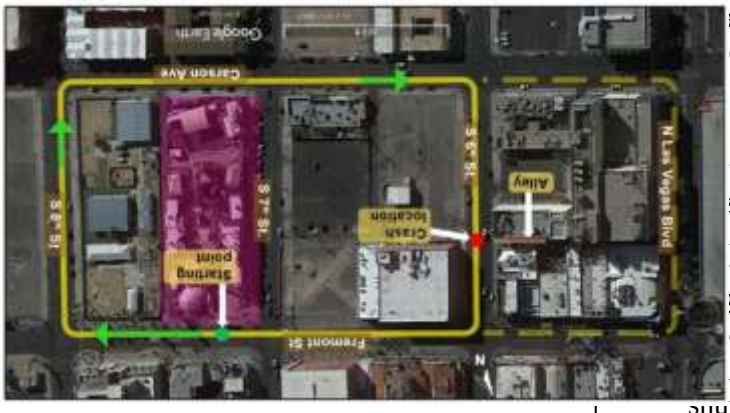


**FIGURE 6: Designated path, direction of travel and location of alley. North up.**

The truck driver saw the pedestrian clear to the left of the truck and started backing up. The attendant and the passengers became concerned that the tractor-trailer was on a trajectory to collide with the shuttle, and waved to get the driver's attention. Four cameras inside the shuttle showed the attendant and passengers waving to the truck driver.

The tractor-trailer driver continued in reverse. While backing up tractor-trailer driver turned his attention to the right, which was when the truck hit the shuttle.

The shuttle attendant believed the shuttle was visible to the truck driver in the right-side mirror from the time the shuttle stopped until the collision. In a post-accident analysis of sight angles, parts of the surrogate shuttle were visible through the windows and in the mirrors on the right side of the tractor-trailer.

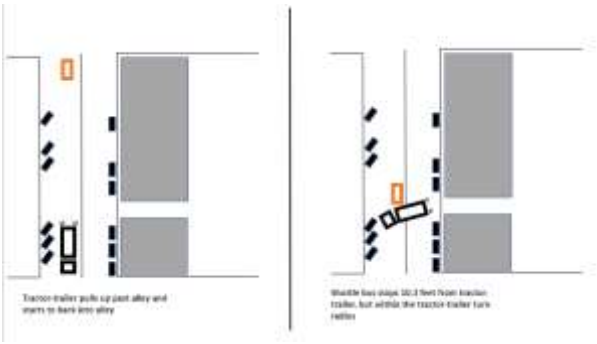

FIGURE 7: Sequence of events leading to collision. Shuttle bus stops short of tractor-trailer but within the turn radius of the reversing tractor-trailer

Eleven seconds after the shuttle stopped, according to the incident report, it was struck by the right front tire of the slow-moving tractor-trailer.

The attendant said that he considered switching to manual mode to move the shuttle, but that he had very little time. He further stated that manual mode was not designed or intended to be used as an emergency mode. That statement was consistent with shuttle operator's policy, as reported to NTSB investigators.

## Discussion

The National Transportation Safety Board (NTSB) determined that the "probable cause" of the collision between the tractor-trailer and the autonomously operated shuttle was:

"the truck driver's action of backing into an alley, *and his expectation that the shuttle would stop at a sufficient distance from his vehicle to allow him to complete his backup maneuver*. [10].

This is indicative of a missing situation in the design and testing of the G&CS. Given the distance at which the shuttle bus stopped from the tractor-trailer, perhaps the design did not adequately distinguish between a tractor-trailer crossing the street and a tractor-trailer backing-up with an associated turn radius.

Likewise in testing, the G&CS was not exposed to this "corner-case" in such a way as to confirm the ability of the system to address this scenario.

The NTSB also cited as "contributing to the cause" of the collision was the:

"attendant not being in a position to take manual control of the vehicle in an emergency."[10].

These conclusions provide three issues for the design, testing and certification of autonomous guidance and control systems:

1. Need to generate a complete and comprehensive set of training data including "Corner-cases" with sufficient repetition of Machine Learning Supervised Training
2. Need to generate a complete and comprehensive set of training data including "Corner-cases" with sufficient repetition of Machine Learning Supervised Training
3. Need to design a user-interface to support the task assigned to the "attendant" or "safety driver" Tasks include monitoring but also meaningful and timely intervention.

These issues and proposed solutions are summarized in Table 1 and discussed below.

**TABLE 1: Three issues and proposed solutions to address the challenges in design, testing and deployment of guidance and control systems for autonomous vehicles**

| Challenges in Design, | Issue | Proposed Solutions |
|---|---|---|

| Testing and Deployment | | |
|---|---|---|
| Design | Generating a complete and comprehensive set of training data including “Corner-cases" with sufficient repetition for Machine Learning Supervised Training | Fast-time Emergent Scenario Simulation (FTESS) |
| Testing & Demonstration of Safety Compliance | Generating a complete and comprehensive set of training data including “Corner-cases" to ensure complete and comprehensive testing | Scenario Generation for Augmented Reality Testing of Vehicle on a Treadmill. Note: Continues simultaneously with deployment to identify potentially hazardous scenarios before they occur in the field |
| Performance of Safety Driver for Testing Period in Responding to High Risk Events during Testing | Absence of explicit design of the procedures, tasks and user-interface to support the task | Safety Driver Monitoring Systems (SDMS) |

### *Generating Sufficient Number of Corner-Cases for Supervised Training*

The accident on the first loop of the first day of operations, demonstrates that the real-world urban environment is complex beyond the imagination of the designers. Even with the best engineering practices and a simple route, there are nuanced circumstances that can arise that are not appropriately captured in the G&CS.

In this case, the G&CS correctly identified a tractor-trailer obstruction. The G&CS brought the shuttle to a stop 10.3 feet from the tractor-trailer. This distance was appropriate for a tractor-trailer crossing the street, but not for a tractor-trailer backing into an alley *with a turn-radius*. The headway distance assigned in the G&CS did not account for the trajectory resulting from a turning radius of the tractor-trailer.

When the environment in which the G&CS operates is relatively finite and simple, traditional engineering practices can account for all the possible combinations of situations that can occur. One example relatively finite and simple environment is aircraft that operate in three-dimensional airspace that is relatively sterile compared to an urban street.

When the environment in which the G&CS must operate is complex and/or infinite, traditional engineering practices are limited. In these cases, rather than designing the G&CS behavior using humans imagining the situations and corresponding behavior, the engineering is conducted by Machine Learning algorithms that record, process, and encode the behavior of expert human operators. The ML algorithms must be exposed to *all* scenarios with sufficient frequency that the ML can: (i) encode all plausible situation-behavior pairs, and (ii) encode subtle differences in situations that require completely different responses from the G&CS.

This approach involves logging millions of miles of real-world driving to attempt to capture all plausible situation-behavior pairs. For example, if situations associated with fatal accidents are taken as the least likely to occur and these accidents occur with 1E-8 per mile, then autonomous vehicles need to perform at least 1E-10 per mile to get exposed to these situation-behaviors (i.e. 100 times more miles). The statistical "rule of 3" says that if N data points with a specified event are to be observed, then the 95% upper bound on the event estimate is 3/N. To achieve 3/N, that would require 3E-10 miles. If a manufacturer has 1E-7 miles driven to date, then they need to repeat the testing completed to date 3,000 times (with no fatalities).

Sherry, Shortle, Donohue, Donnelly [11] and Nanduri & Sherry [12] proposed a “Fast-Time Emergent Scenario Simulation (FTESS).” FTESS is an agent-based, rare-event simulation “digital-twin” of the system.

FTESS can be used to expand the training data for the ML algorithms with significant reduction is time. The FTESS starts with a "seed" scenario from the real-world driving and using Monte Carlo techniques on a super-computer generates variances on the seed scenario.

The FTESS leverages techniques for rare-event simulation, edge computing, and runs on a super-computer. There are two main approaches for improving the efficiency of rare-event simulations— importance sampling (IS) and splitting [13], [14], [15]. The idea of IS is to change the underlying sampling distribution so that rare events are more likely. The idea of splitting is to create separate copies of the simulation whenever the simulation gets ''close'' to the rare event of interest, effectively multiplying promising runs that are more likely to reach the rare event. Splitting is useful for systems that tend to take many incremental steps on the path to the rare event. IS is also useful for systems that tend to take a small number of ''catastrophic jumps'' to the rare event.

Inexpensive access to cloud-based Super-computing. GMU has access to the Argos Supercomputing Cluster. Ph.D. students are trained to run parallel computing algorithms on the cluster.

Agent-based models. MASON is a fast discrete-event multiagent simulation library core in Java, designed to be the foundation for large custom-purpose Java simulations, and also to provide more than enough functionality for many lightweight simulation needs. MASON contains both a model library and an optional suite of visualization tools in 2D and 3D. Ph.D. students are trained to develop agent-based models in MASON (Calderon-Meza, Sherry, 2009).

The FTESS can exhibit varying levels of model fidelity according to the development phase

### *Generating a Complete and Comprehensive set of Scenarios to Ensure Safety Compliance*

The use of Machine Learning in the design of the G&CS results in "black-box" automaton that prohibits the utilization of regulator approved traditional testing methods. The Machine Learning G&CS cannot undergo a code-inspection or any form of input-output testing as the input-output relationship is "hidden" in the machine learning code algorithm.

The alternative means of compliance is via performance/risk-based testing in which the system performance is recorded during an extended testing by demonstration. The challenge, however, is the number of miles that is required to demonstrate safety to the required target level of safety may only be completed in several years, sometimes after the technology is obsolete.

For example, if the US driving fatality rate is 1E-8 per mile, then autonomous vehicles need to be 100 times safer (i.e. 1E-10 per mile). The statistical "rule of 3" says that if N data points are observed with zero fatalities, then the 95% upper bound on the fatality estimate is 3/N. So that would require 3E-10 miles with zero fatalities to demonstrate 1E-10 per mile. If a manufacturer has 1E-7 miles driven to date, then they need to repeat the testing completed to date 3,000 times (with no fatalities).

An alternative approach is to create an Autonomous Vehicle test-bed that is akin to what is known in the aviation industry as an "iron-bird." A complete aircraft with a full set of actuators that are commanded by flight control automation that responds to artificial sensor inputs generated by a simulator. In the case of a shuttle bus, the Autonomous Vehicle Certification Test-bed will have the car on an automobile-treadmill. The LIDAR, Radar, camera, GPS, etc. sensors will be disconnected and the shuttle bus guidance and control automation will be fed by data from a simulator. The simulator data will include LIDAR maps, Radar profiles, digital images etc of real-world scenarios. The vehicle will need to demonstrate the correct response at each point in the unfolding scenario.

The hard part is generating the scenarios. The proposal for AV Certification is to collect the scenarios from real-world driving with a human operator. The scenarios would be ranked based on their risk with the higher risk scenarios being tested before the lower risk scenarios. A bank/data-base of AV test-cases would be created and would be a continuous work in progress adding new high risk scenarios as they are uncovered in the real world.

These test-cases would be supplemented by additional scenarios and variations on the scenarios

through the "Fast-Time Emergent Scenario Simulation (FTESS) described above.

Another source of scenarios is voluntary operator reports (e.g. ASRS, ATSAP). The scenarios described in these reports will serve as the "seed" from which the FTESS can generate additional variations of the scenarios. In this way the FTESS can strive towards an exhaustive search of the hazard-space before the voluntary report becomes an accident.

Yet another source of scenarios is the National Automotive Sampling System (NASS). The NASS provides the National Highway Traffic Safety Administration an efficient and and reusable resource with which to conduct data collection. The Data in NASS focus on passenger vehicle crashes and could be used to generate scenarios for machine learning[16]

Although the FTESS may never generate all the plausible combinatorics, it will generate more than can be tested using traditional "test-case" approaches.

Once the system is deployed, the safety analysis not over. The Autonomous Vehicle Certification Test-bed is used while the systems are fielded. The FTESS can be run in fast-time shadow mode, using the actual operations as the starting point of the fast-time simulation (e.g. aircraft crossing the final approach fix).

### *Explicit Design of Attendant Tasks*

When an operator is inserted to monitor and intervene in the event in an inappropriate G&CS command it cannot be as an afterthought. To meaningfully support this task, the design of autonomous G&CS must be explicitly designed to support the task. For example, the "attendant" in the shuttle bus incident was not supported by automation to identify situations and their appropriate G&CS behavior. Further the procedures for intervention were absent. For example, the attendant was required to perform a task on a hand-held controller that was located in a glove-box. Further the input devices on the controller (e.g. buttons, joysticks) were not labelled and had similar buttons with opposite effects.

The operator must have specified monitoring and intervention procedures. All airliners are certified with emergency procedures that are tested extensively. Further the pilots are required to demonstrate proficiency in these emergency procedures approximately once every 12 months.

The G&CS automation must be explicitly designed to support human the monitoring and intervention procedures. Specifically, the G&CS must annunciate:

1. the intent of the G&CS (e.g. emergency braking, stopping constant speed, acceleration to constant speed, …).
2. the emerging situation perceived by the G&CS (e.g. traffic light green, vehicle ahead slowing, …).

If the G&CS perceived situation and/or intent do not match with the operator's perception, the operator can make an informed decision to intervene.

The status of the autonomous G&CS must be clearly annunciated at all times so there is no ambiguity if the attendant is in manual control

In addition to a fail-safe emergency stop function, there should also be a means to command the vehicle to move from a stopped position to safer location. This repositioning must be accomplished in the available operational time window (AOTW) appropriate for the circumstances (e.g. less than 10 seconds)

The operator should also have a means to communicate to the outside world to draw attention to the location of the vehicle (e.g. flashing lights, horn, loud-speaker)

**Safety Driver Monitoring System**

This section outlines the design of user-interface to support the attendant.

Since the Machine Learning G&CS for Autonomous Shuttle Buses cannot be certified using traditional "code inspection & I/O test methods used in the avionics industry, Autonomous Shuttle Buses must drive millions of miles and demonstrate their safety. This approach is known as "Risk-based" or "Performance-based" Certification.

The challenge with "Performance-based" Certification is that it requires a "Safety Driver"

(also known as the “attendant”) to monitor the emerging situations in which the Autonomous Shuttle Bus is operating and the autonomous G&CS response.

In this way the "Safety Driver" really becomes a "Safety Passenger" that has to use mental discipline to remain engaged in the process of driving when during testing approximately 99.999% of the time the G&CS works as expected.

The challenge is to engage the human operator such they remain a "safety driver" and do not become a "safety passenger."

The Safety Driver Monitoring System (SDMS) provides the human operator a display with the "situation" that the guidance and control system "sees." The safety driver must then conform or reject the assessment made by the automation. For example, the guidance and control system can "see" a pedestrian crossing the road between intersections and display Situation: "Pedestrian crossing left to right, 400 ft." The operator would hit a button "Correct" or "Not Correct"

The system could also display the G&CS system response e.g. Behavior: "Slowing to a stop 10 feet from Pedestrian" The operator would hit a button "Appropriate" or "Not Appropriate"

There are several challenges with this proposal.

1) The safety driver would have a tremendous problem with divided attention and working memory overload. An experienced driver will be making most of the judgments needed while driving automatically without conscious awareness. The proposed monitoring task would force the drivers to attend to the autonomous system while their attention would likely be elsewhere and the response would also interfere with the driving task. For example, the driver might see the pedestrian crossing, non-consciously process the speed and decide that the pedestrian is not a factor but the vehicle slowing quickly in front is. The system could ask about the pedestrian at that point. The driver does not have access to that memory, so s/he scans the environment to locate the interloper and takes his/her attention away from the actual threat.

2) The safety driver would become a research assistant. If the autonomous system is good enough to be deployed, it is likely that the safety driver would prioritize responding to the system rather monitoring the environment. Hence, the task that you want the driver to perform -- scanning for undetected threats -- would not be routinely performed.

## Conclusions

The analysis of the Tractor-Trailer/Shuttle Bus incident provides anecdotal evidence of the challenges in designing, testing and certifying autonomous vehicles. The analysis identified three main issues:

1. the analysis demonstrates the difficulty in designing for all the “corner-cases” for safe fielding of an autonomous system.
2. the analysis demonstrates the difficulty in demonstrating safety compliance to a target level of safety for systems developed using machine learning that cannot be tested using traditional testing methods such as code-inspection or forms of input-output testing.
3. the task of the "safety driver” used during the extensive testing phase to prevent hazardous events must be an explicit, intentional design, not an ***afterthought*** with appropriate user-interfaces for monitoring and intervening in rare-events.

The paper suggests leveraging a “digital twin” simulation of the environment to generate scenarios that can be used for supervised training of the machine-learning guidance and control system, and for augmented reality testing environment.

The paper also describes a user-interface to support an assisted driver.

## References

[1] Draper, C. S.; Wrigley, W.; Hoag, G.; Battin, R. H.; Miller, E.; Koso, A.; Hopkins, A. L.; Vander Velde, W. E. (June 1965). Apollo Guidance and Navigation (PDF) (Report). Massachusetts: Massachusetts Institute of Technology, Instrumentation Laboratory. pp. I-3 et seqq. Retrieved October 12, 2014.

[2] McRuer D., Ashkenas I., Graham D. (1973) Aircraft dynamics and automatic control. Princeton University Press, Princeton, NJ

[3] Sherry (1995) A Formalism for the Specification of Operationally Embedded Reactive Systems. Volume5, Issue1. St Louis, MO, July 22–26, 1995. Pages 571-578.

[4] Guo, X., L. F. Neale, M. Westcott (2009) Target Level of Safety Measures in Air Transportation – Review, Validation and Recommendations, In Proceedings The IASTED International Conference on Management Science and Risk Assessment (AMSRA 2009) & Modelling, Simulation and Identification (MSI 2009) Beijing, China.

[5] NHTSA (2019) Federal Motor Vehicle Safety Standards (FMVSS). Available at https://www.nhtsa.gov/laws-regulations/fmvss

[6] Koopman, P. U. Ferrell, ,F. Fratrik, M. Wagner. (2019) A Safety Standard Approach for Fully Autonomous Vehicles. WAISE 2019. https://users.ece.cmu.edu/~koopman/pubs/Koopman19_WAISE_UL4600.pdf

[7] Kalra, N., S. M.Paddock (2016) Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability? Transportation Research Part A: Policy and Practice. Volume 94, December 2016, Pages 182-193

[8] Wei, J., J. M. Snider, J. Kim, J. M. Dolan, R. Rajkumar and B. Litkouhi (2013) Towards a Viable Autonomous Driving Research Platform. 2013 IEEE Intelligent Vehicles Symposium (IV). June 23-26, 2013, Gold Coast, Australia

[9] Reschka A. (2016) Safety Concept for Autonomous Vehicles. In: Maurer M., Gerdes J., Lenz B., Winner H. (eds) Autonomous Driving. Springer, Berlin, Heidelberg.

[10] National Transportation Safety Board (2019) Highway Accident Brief Low-Speed Collision Between Truck-Tractor and Autonomous Shuttle, Las Vegas, Nevada, November 8, 2017. Accident Number: HWY18FH001. Accident Type: Collision involving automated test vehicle on public road. National Transportation Safety Board Washington, DC 20594. https://www.ntsb.gov/investigations/AccidentReports/Reports/HAB1906.pdf

[11] Sherry, L. J. Shortle, G Donohue, O. Donnelly (2019) Fast-Time Emergent Scenario Simulation (FTESS). Report: Center for Air Transportation Systems Research at George Mason University. Report # CATSR 2019-11.

[12] Nanduri, A., L. Sherry (2016) Generating Flight Operations Quality Assurance (FOQA) Data from the X-Plane Simulation. In Proceedings 2016 Integrated Communications, Navigation, Surveillance (ICNS) Conference, Dulles, Va. April 19-21, 2016

[13] J. F. Shortle and P. L'Ecuyer, "Introduction to Rare-Event Simulation," in *Wiley Encyclopedia of Operations Research and Management Science*, Hoboken, NJ, USA: John Wiley & Sons, Inc., 2011, p. eorms0006.

[14] Zare-Noghabi, A., & Shortle, J. (2017). Rare event simulation for potential wake encounters. 2017 Winter Simulation Conference (WSC), 2554–2565. https://doi.org/10.1109/WSC.2017.8247983

[15] Snisarevska, O. L. Sherry, J. Shortle, G. Donohue (2018.) Balancing Throughput and Safety: An Autonomous Approach and Landing System, In Proceedings IEEE ICNS Conference 2018. April, 2018

[16] "National Automotive Sampling System (NASS) | NHTSA." [Online]. Available: https://www.nhtsa.gov/research-data/national-automotive-sampling-system-nass. [Accessed: 16-Feb-2020].

## Acknowledgements

The authors acknowledge the contributions of Jonathan West, Seungwon Noh, Jomana Bashata, Sasha Donnelly (Center for Air Transportation Systems Research at George Mason University)..

## Email Addresses

lsherry@gmu.edu