A FRAMEWORK FOR AUTOMATED DRIVING SYSTEM TESTABLE CASES AND SCENARIOS

Paul Rau
National Highway Traffic Safety Administration
United States

Eric Thorn
Southwest Research Institute
United States

Paper 19-0301

ABSTRACT

Automated Driving Systems (ADS) are being developed to perform the primary functions of the dynamic driving task (DDT). These technologies hold great promise to improve safety and mobility for transportation. Test scenarios are critical for assessing the safety assurance of ADS in a range of operational environments and roadway conditions. The development of testing scenarios for ADS is proving to be an important challenge for the development of safety assurance requirements, certification and licensing frameworks, testbed services, standards, and international harmonization.

This paper summarizes foundational research undertaken to identify a sample preliminary, objective testing and evaluation approach for ADS. The paper considers technologies of interest that fall within Level 3 through Level 5 of the SAE International levels of driving automation and identifies a cross-section of prototype and conceptual ADS that are then categorized into seven generic ADS features.

This research also takes the first steps to partition the ADS performance space by identifying and assessing the primary variables that comprise an ADS test scenario. Those primary variables are described in detail, and include:

- Tactical and Operational Maneuvers
- Operational Design Domain (ODD)
- Object and Event Detection and Responses (OEDR)
- Failure Mode Behaviors

Tactical and operational maneuver capabilities largely focus on the control-related elements of the DDT (i.e., lateral and longitudinal control) that enable an ADS to navigate to reach its destination (e.g., lane centering / following, turning). A working list of these capabilities is presented. The ODD represents the operating conditions under which an ADS is designed to function (e.g., roadway types, weather conditions, etc.). A notional hierarchical ODD taxonomy is presented and described. OEDR capabilities include the elements of the DDT that involve monitoring the driving environment and implementing appropriate responses to relevant objects and events. A working list of OEDR capabilities is presented. Failure mode behaviors include fail-safe (FS) and fail-operational (FO) strategies that will allow an ADS to respond to a variety of failures, including DDT performance-relevant system failures that require the ADS or a DDT fallback-ready user to achieve a minimal risk condition.

The paper also considers the implementation of the proposed evaluation framework using existing test methods, including modeling and simulation (M&S), closed track testing, and open road testing. It further seeks to examine how each of the testing methods can be logically used to minimize the complexity of comprehensive safety assessments of ADS by leveraging each method's strengths to maximize the knowledge gained from each test. It also includes extensive discussion of challenges associated with testing ADS, including challenges related to the technology itself as well as challenges associated with test execution. This paper is based on research completed by NHTSA and its contractors, and is more fully documented in NHTSA Report DOT HS 812 623, "A Framework for Automated Driving System and Testable Cases and Scenarios"; September 2018.

INTRODUCTION

Since 1975, the first year that the Fatality Analysis Reporting System began collecting data, the rate of traffic fatalities per 100 million miles traveled in the United States has decreased by 66 percent, according to the National

Highway Traffic Safety Administration’s (NHTSA’s) Traffic Safety Facts 2015 data (NHTSA, 2017b). Advancements in motor vehicle safety have been made through continuous engineering innovation, public education, industry agreements, safety regulations, and safety rating programs. There is, however, significant room for continued focus on motor vehicle traffic safety. In October 2017, NHTSA reported that traffic fatalities increased by 5.4 percent from 2015 to 2016 (35,485 to 37,461) for the United States (NHTSA, 2017c), which follows an 8.4 percent increase from 2014 to 2015 (32,744 to 35,485) (NHTSA, 2017b).

Many forces are at work in the automotive industry to advance safety technology. The worldwide automotive industry has recognized driver performance (e.g., error and choice) as a key factor that impacts safety and has begun to introduce systems that complement the driver in terms of enhanced perception with 360-degree vehicle views and rear video systems. Advanced Driver Assistance Systems that monitor the operational environment and enhance driver detection and response, such as Forward Collision Warning (FCW) and Lane Keeping Assist (LKA), are increasingly common in newer model vehicles. Additionally, 20 automakers have committed to making Automatic Emergency Braking (AEB) a standard feature in new vehicles by 2022 (IIHS, 2016).

Recently, research activities by several companies to develop ADS that can perform certain driving functions automatically have captured the nation's attention. ADS have been the subject of multiple congressional hearings and the public has provided numerous responses to NHTSA’s Federal Automated Vehicles Policy (NHTSA, 2016b), including over 1,100 responses from industry participants, state and municipal transportation agencies, policy groups, and citizens (Kyrouz, 2017). The United States Department of Transportation (USDOT) and NHTSA recently released an update to their federal guidance for ADS that focused on their development and safe deployment and operation. NHTSA also continues to advance its ADS research. The research summarized in this paper sought to analyze aspects of ADS testing and develop examples of tests and evaluation methods for specific ADS features. A sample testing framework was developed that could further support the goals of improving safety for all users of the transportation network.

OBJECTIVE

The purpose of this study was to analyze aspects of ADS testing to create a framework for developing test cases and test scenarios for ADS. Consideration was given to keeping the framework flexible and extensible such that it could be applied with different test approaches and methods. The ultimate goal of this framework is to support the safe deployment of ADS in the broader transportation system.

AUTOMATED DRIVING SYSTEM FEATURES

As an initial step to develop this framework, sample concept ADS features that have been proposed for deployment were identified. This analysis focused on SAE International Levels 3-5 ADS (SAE International, 2018), such as Google’s (Waymo’s) self-driving car project and others like it that focus on next-generation automation. This step is critical because the sample concept ADS features are used to identify ODDs and OEDRs, develop preliminary tests and/or evaluation methods, and assess FS and FO mechanisms, which form a foundation to begin considering validation and verification approaches for ADS.

A four-stage approach was followed to identify ADS features: 1) review the literature, 2) define a framework for discussing ADS features, 3) define features and behaviors, and 4) categorize the features. To guide later analysis, priority ADS features on which to focus were identified. Over 50 literature sources were reviewed, including original equipment manufacturer (OEM) websites, press releases of vehicles being tested in specific domains, NHTSA pre-crash scenario analysis (NHTSA, 2007), NHTSA’s Fiscal Year 2017 budget request (NHTSA, 2016c), NHTSA L2 and L3 Human Factors Concepts (NHTSA, 2015), Federal Highway Administration (FHWA) managed lane use cases (FHWA, 2008), and technical and international publications, including proceedings of the 2015 and 2016 Automated Vehicles Symposiums and United Nations Economic Commission for Europe (UNECE) World Forum for Harmonization of Vehicle Regulations (WP.29) Automatically Commanded Steering Function working group, among others. Research sponsored by USDOT, such as the Crash Avoidance Metrics Partnership Automated Vehicle Research for Enhanced Safety (Christensen, et al., 2015; NHTSA, 2016d), which details functional descriptions for on-road driving automation levels, was also used.

Twenty-three concept ADS features were identified:

1. Audi Traffic Jam Pilot

2. Audi Highway Pilot
3. Auro Self Driving Shuttle
4. Baidu Automated TNC[1] Bosch Valet Parking
5. CityMobil2 Automated Shuttle
6. Bosch Highway Pilot
7. EZ10 Self Driving Shuttle
8. Ford Automated TNC
9. GM Cruise Automation TNC
10. Google Car
11. Honda Automated Drive
12. Mercedes Highway Pilot Truck
13. Navya Arma Shuttle
14. Nissan Autonomous Drive
15. Olli Local Motors Shuttle
16. Otto Trucking
17. Tesla Self-Drive
18. Toyota Chauffeur
19. Toyota Guardian
20. Uber Automated TNC
21. Varden Labs Self Driving Shuttles
22. Volkswagen I.D. Pilot
23. Volvo IntelliSafe Auto Pilot

These 23 features were categorized into the following seven generic categories:

1. L3 Conditional Automated Traffic Jam Drive
2. L3 Conditional Automated Highway Drive
3. L4 Highly Automated Low Speed Shuttle
4. L4 Highly Automated Valet Parking
5. L4 Highly Automated Emergency Take-Over
6. L4 Highly Automated Highway Drive
7. L4 Highly Automated Vehicle / TNC

Through the literature review and analysis, a working list of tactical and operational maneuvers related to ADS driving control was created. Some examples of these tactical and operational maneuvers included: parking, maintaining speed, lane centering, low-speed merge, right-of-way decision, following driving laws, and U-turns, among others.

Each of the identified generic ADS features was then described in terms of tactical maneuver behaviors, estimated commercial availability, and estimated level of automation. Figure 1 shows a sample analysis for the L4 Highly Automated Vehicle / TNC Feature. It should be noted that these commercial ADS features were identified several years ago and the list has changed quite significantly. It should also be noted that the estimates for commercial availability, level of driving automation, and tactical maneuver demonstration were deduced from the information available, which was limited, and as such should be considered notional.

[1] TNC: Transportation Network Company

ADS Features and Tactical Maneuvers (X = demonstrated, ? = speculated)	**Commercially Available? (Y/N)**	**Level of Automation (SAE 1-5)**	Parking	Maintain Speed	Car Following	Lane Centering	Lane Switching/Overtaking	Enhancing Conspicuity	Merge	Navigate On/Off Ramps	Follow Driving Laws	Navigate Roundabouts	Navigate Intersection	Navigate Crosswalk	Navigate Work Zone	N-Point Turn	U-Turn	Route Planning
Waymo Automated TNC	N	4	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Tesla Self-Drive	N	4	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Volkswagen I.D. Pilot	N	4?	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Volvo IntelliSafe Auto Pilot	N	4	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Nissan Autonomous Drive *(2020)*	N	4?	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
GM Cruise Automation	N	4	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Uber Automated TNC	N	4	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Honda Automated Drive *(2020)*	N	4	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Ford Automated TNC *(2022)*	N	4	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Baidu Automated TNC	N	4	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Toyota Chauffeur	N	4	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Figure 1. L4 Highly Automated Vehicle / TNC Features.

OPERATIONAL DESIGN DOMAIN

An operational design domain (ODD) describes the specific operating domain(s) in which an ADS feature is designed to function with respect to roadway types, speed range, lighting conditions (day and/or night), weather conditions, and other operations constraints. ODD will likely vary for each ADS feature, even if there is more than one ADS feature on a vehicle. The testing framework presented in this paper considers the potential range of ODDs and how ODDs factor into developing potential test cases.

A three-stage approach was taken to define the ODDs:

1. Review the literature, including popular media, press releases, technical journals, and conference proceedings to identify key concepts, enumerate potential ODD characteristics, and examine approaches to ODD in other industries.
2. Define and categorize ODD into a sample taxonomy that can be used by departments of transportation (DOTs) and industry to discuss ADS.
3. Describe ODDs in which concept ADS features may operate based on literature review and engineering judgment.

Over 50 literature sources were reviewed, including OEM websites, press releases, USDOT documents, including NHTSA pre-crash scenario analysis and FHWA managed lane use case, as well as technical and international publications, including proceedings of the 2015 and 2016 Automated Vehicles Symposiums. Additionally, the NHTSA Fiscal Year 2017 Budget Request to Congressional Appropriations Committees (NHTSA, 2016c) identifies several ADS use cases that were considered when defining the ODD for this analysis. It should be noted that given the emerging and highly competitive nature of ADS technology, it is inherently difficult to obtain explicit and complete information about the intended ODD of an ADS feature. In the absence of information about an ODD, engineering judgement was used at times to define the ODD taxonomy and identify the ODD for concept ADS features.

While the literature provided many examples of ODD elements, no classification framework was identified. This work takes an initial step towards developing a taxonomy to organize the many ODD elements identified in research. This sample ODD taxonomy takes the form of a hierarchy of categories and subcategories, each with definitions and, where appropriate, gradations. This taxonomy is meant to be descriptive, not normative, as it is envisioned that these elements may be organized into several different groupings. The taxonomy offers a structured approach to organize and identify various ODDs for ADS features, especially when there are several different possible combinations. Figure 2**Error! Reference source not found.** shows the broad range of top-level categories and immediate subcategories. It should again be noted that this sample taxonomy was derived using available information at the time the research was conducted and should be considered notional.

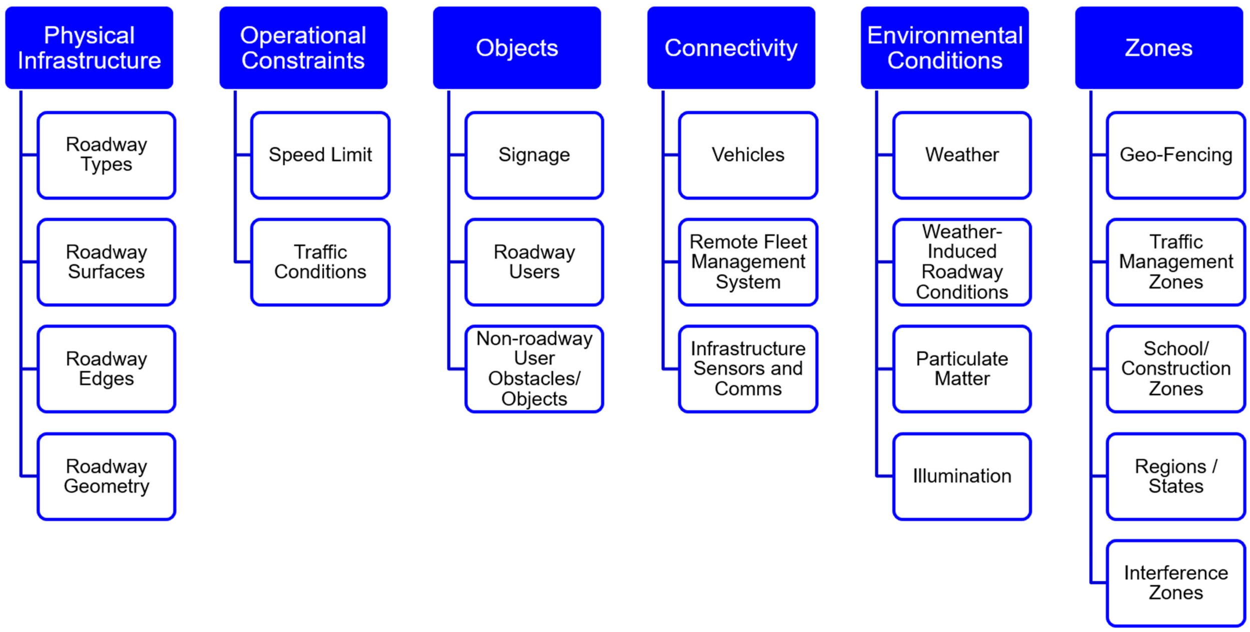

Figure 2. ODD Classification Framework with Top-Level Categories and Immediate Subcategories.

The hierarchy extends into multiple sublevels. For example, **Error! Reference source not found.** shows that the "Environmental Conditions" category was divided into four subcategories: weather, illumination, particulate matter, and road weather. Weather is further subdivided into rain, temperature, wind, and snow. For this research, it was helpful to further subdivide rain into gradations to capture the data that were collected on ADS features. For example, some ADS features had been tested in light rain, while others had been tested in heavy rain. Although the application of this sample taxonomy has been useful in the context of this research, further research and stakeholder engagement would be beneficial in refining and objectively quantifying the categories and gradations.

The sample ODD taxonomy lends itself to serving as a checklist for identifying the ODD of an ADS feature. A comprehensive ODD checklist was generated based on the ODD taxonomy described above. To demonstrate a potential application of the checklist, the checklist was filled out for three theoretical ADS features. It should be noted that the manufacturer determines the ODD for a feature, and the ODD may vary for similar ADS features. The theoretical features presented here are purely demonstrative, not representative of any commercially marketed ADS feature.

To test a vehicle's ability to operate safely, ODD is considered in test development and execution. Scenarios consider a combination of ODD elements that can be used to describe conditions for test cases and scenarios; for example, a highway with a concrete surface with a light mist. Test facilities are limited in their ability to re-create certain ODDs (e.g., urban environments, hill crests) and may need to be upgraded with new infrastructure to support testing. Some ODD elements are difficult to quantify and re-create (e.g., weather), though it is possible they could be addressed through functional safety design practices and on-road testing.

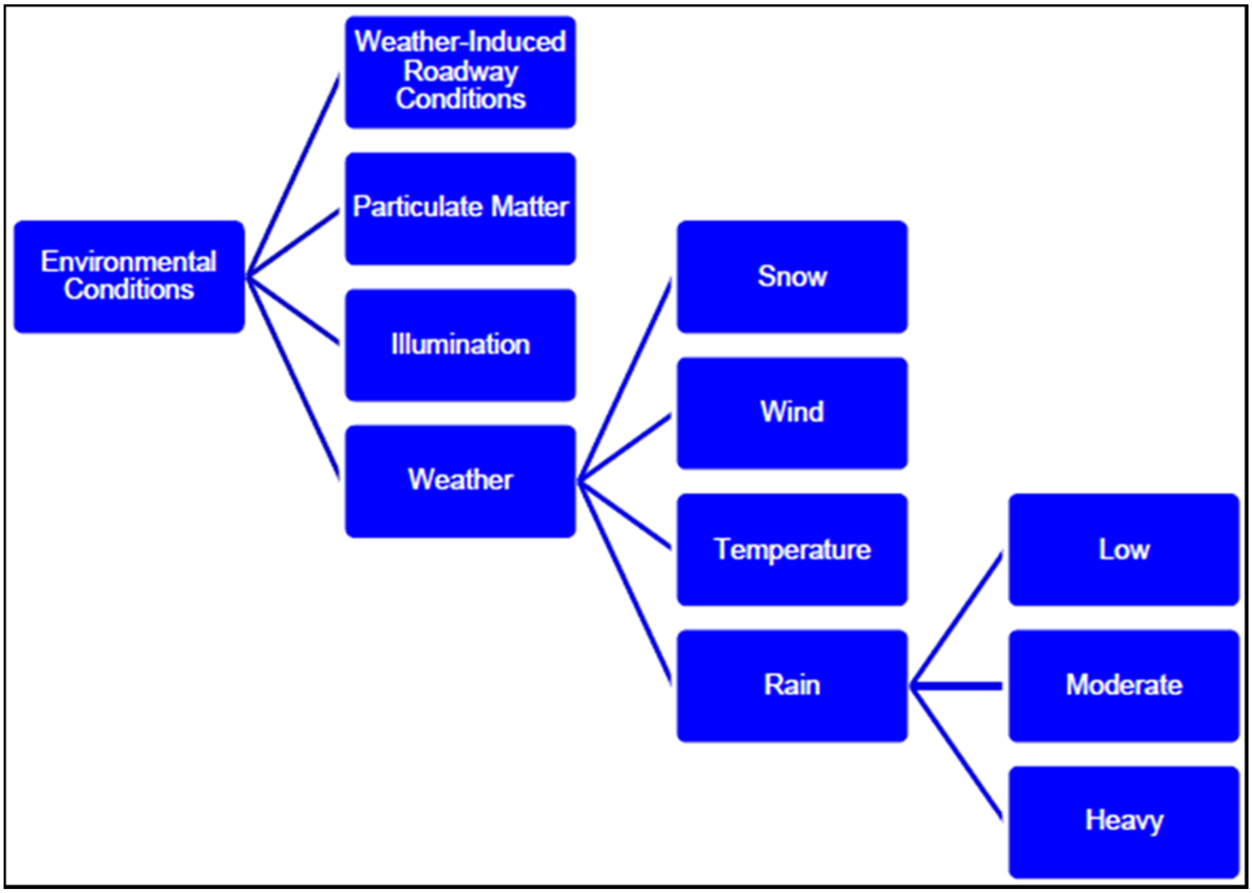

Figure 3. Example of Hierarchy Levels within the Environmental Conditions Category.

OBJECT AND EVENT DETECTION AND RESPONSE

While performing tactical maneuver behaviors described previously, ADS will inevitably interact with a variety of static and dynamic physical objects that may alter how these behaviors are executed. SAE J3016 identifies the following real-time functions as elements of the DDT related to addressing these interactions with objects:

- Object and event detection, recognition, and classification
- Object and event response

These functions can be generalized under the term Object and Event Detection and Response (OEDR). OEDR represents the ability of the ADS feature to detect any circumstance that is immediately relevant to the driving task and implement an appropriate response. One of the factors that determines the level of driving automation of an ADS is whether the human driver or ADS is responsible for monitoring the driving environment. ADS, which were the focus of this research, range from SAE International L3 through L5, which means that the ADS feature is completing all aspects of monitoring the driving environment.

The elements of an ADS functional architecture that are specifically relevant to OEDR generally include hardware and software components that support:

- Sensing (e.g., radar, laser scanners, cameras, etc.)
- Perception (e.g., road feature classification, object segmentation and classification, etc.)

- World modeling (e.g., persistent data mapping, dynamic obstacle tracking, and prediction, etc.)
- Navigation and planning (e.g., path planning and motion control commands to implement responses)

The sensing and perception elements of the architecture specifically support detection of relevant objects. World modeling supports the aggregation of perception and other information to identify and understand events that may occur through interactions with those objects. Navigation and planning support determination of the appropriate response to those events and interactions, and the generation of control commands to implement that response.

Three of the generic ADS features were selected for an OEDR analysis (L3 Conditional Automated Traffic Jam Drive, L3 Conditional Automated Highway Drive, and L4 Highly Automated Vehicle / TNC). This allowed for an evaluation of a cross-section of operating environments and conditions, as well as driving scenarios. Following the evaluation of the operational needs of the selected ADS features, a focusing exercise established baseline ODDs for each feature to further refine the analysis to identify OEDR capabilities for the three selected features. This exercise served to frame the OEDR analysis to account for the potential variability of certain ODD elements, as well as the substantial number of combinations and permutations of ODD elements. It is reasonable to expect that different organizations developing similar ADS features will generate unique designs and implementations, and thus will ultimately define different ODDs for their respective systems. With the ODD baselines established for each feature, a survey and analysis of the driving scenarios resulting from the operations descriptions led to the identification of relevant objects and interactions that the ADS could encounter. These objects and events are derived from an evaluation of normal driving scenarios for a given ADS feature operating in its ODD.

The developed baseline ODDs were used to identify important objects and events that ADS could feasibly encounter within those ODDs. Aggregated OEDR behavior capabilities are shown in Table 1.

Table 1. Summary of OEDR Behavior Capabilities.

Detect & Respond to Speed Limit Changes	Detect & Respond to Relevant School Buses
Detect & Respond to Encroaching, Oncoming Vehicles	Detect & Respond to Relevant Emergency Vehicles
Perform Vehicle Following	Detect & Respond to Relevant Pedestrians
Detect & Respond to Relevant Stopped Vehicles	Detect & Respond to Relevant Pedalcyclists
Detect & Respond to Relevant Lane Changes / Cut-ins	Detect & Respond to Relevant Animals
Detect & Respond to Relevant Static Obstacles in Lane	Detect & Respond to Relevant Vehicle Cut-out / Reveal
Detect & Navigate Work Zones	Detect & Respond to Relevant Vehicle Roadway Entry
Detect & Respond to Relevant Safety Officials	Detect & Respond to Relevant Adjacent Vehicles
Detect & Respond to Relevant Access Restrictions	Detect & Respond to ODD Boundary Transition
Detect & Respond to Relevant Dynamic Traffic Signs	

FAILURE MODE BEHAVIOR

ADS will utilize FO and FS mechanisms when the system does not function as intended. These mechanisms are intended to cause the ADS to attain a minimal risk condition (MRC) that removes the vehicle and its occupants from harm's way, to the best extent possible. Defining, testing, and validating FO and FS strategies for achieving an MRC are important steps in promoting the safe operation and deployment of ADS.

The appropriate failure mitigation strategy and resulting MRC for a given ADS is largely dependent on the type and nature of failures the ADS experiences. To this end, an understanding of potential ADS failure modes is necessary. As such, a high-level failure analysis was performed. The results of this analysis informed the assessment of FO and FS mechanisms. A variety of failure and hazard analysis techniques exist, including fault tree analysis (FTA), system failure mode and effects analysis (FMEA), failure modes, effects, and criticality analysis (FMECA), system-theoretic process analysis, and hazard and operability analysis (HazOp). System FMEA was identified and selected as an initial approach to develop the high-level analysis needed to identify potential failures in each subsystem of the representative functional architecture, as well as their causes and impacts.

Existing reports and literature on ADS failures, including from the Defense Advanced Research Projects Agency (DARPA) Grand and Urban Challenges (DARPA, 2008), as well as engineering judgments and prior experience in ADS development and testing were leveraged and considered. It was assumed that a detailed failure analysis employing a range of techniques noted above has been performed on the base vehicle platform, and therefore efforts were focused on components specifically related to the ADS. This allowed for a deeper dive into a representative ADS functional architecture. Furthermore, failures that could have safety implications, as opposed to failures that are merely an inconvenience, were prioritized. The FMEA was broken down by architecture subsystems to identify potential key failures at each step through the ADS "pipeline":

- Sensing and communication
- Perception
- Navigation and control
- Human Machine Interface (HMI)

In general, many of the ADS failure modes described above could be attributed to some kind of failures by the ADS to obtain information needed to perform the DDT. These were summarized into three primary categories as failures attributed to:

- No data – Information is absent altogether
- Inadequate quality data – Information is of poor or degraded quality
- Latent data – Information is delayed or old

After completing the FMEA for the ADS architecture, the various failure modes and effects were summarized and mapped to the relevant tactical maneuver and OEDR behaviors for the three down-sampled ADS features (L3 Traffic Jam Drive, L3 Highway Drive, and L4 Highly Automated Vehicle/TNC). This notionally provides a mapping from the specific failures identified in the FMEA, to the generalized failures summarized in the previous section, to the behaviors implemented by various ADS features.

Based on the general failure modes identified, potential failure mode responses and strategies were identified. This effort focused on FS strategies for cases where the ADS cannot continue to operate due to a significant failure, and FO strategies for cases where the ADS could continue to operate even in the face of a failure. It should be noted that these potential FS and FO strategies were determined from engineering judgements and available literature, and as such should be considered notional.

The primary goal of an FS strategy is to rapidly achieve an MRC where the vehicle and occupants are safe. Three candidate FS mechanisms were considered for further evaluation:

- Transition vehicle control to fallback-ready user
- Safely stop in lane of travel
- Safely move out of travel lane and stop

FO strategies allow the ADS to continue to function, even in the event of one or more failures. It is important to note that this operation may only be supported for a limited duration, or potentially with a reduced set of capabilities. Three primary FO mechanisms were considered for further analysis:

- Hardware/software redundancy
- Adaptive compensation (e.g., ignore data coming from failed sensor or component and weight inputs from other sensors or components more heavily)
- Degraded operation(s)
    - Reduced top speed
    - Reduced level of automation
    - Reduced ODD
    - Reduced maneuver capabilities
    - Reduced OEDR capabilities

PRELIMINARY TEST AND EVALUTION METHODS

After evaluating prototype ADS features, potential ODDs, potential OEDR capabilities, and potential failure mode strategies, a sample evaluation framework was developed to support the assessment of ADS for safe deployment. Sample test procedures were also developed using engineering judgements, previous test procedure development experience, and use cases. The test framework and procedures developed gave special consideration to achieving repeatability, reliability, and practicality. Lastly, many challenges associated with testing ADS and further research needed to help address these challenges were identified. Challenges included those related to the technology itself as well as test execution.

To identify appropriate methods to evaluate ADS, a review and assessment of existing testing methods and tools was performed. This evaluation served to develop an understanding of how testing is currently being executed for vehicles capable of various levels of automation. It also served to identify potential gaps in this existing testing framework, which led to the identification of additional and modified tools and methods to fill those gaps and helped create a testing framework. This assessment included a meeting with crash avoidance test engineers at NHTSA's Vehicle Research and Testing Center (VRTC) in Ohio to discuss their current testing of vehicles capable of SAE International L1 and L2 driving automation. Findings from the previous analyses were presented and initial thoughts on the steps to develop a useful set of test methods and actual tests were provided.

A common test scenario framework that could be used broadly across the various testing methods and tools was then established. This framework built upon the findings of the previous tasks to include the principal elements of ADS operation (tactical maneuver, ODD, OEDR, and failure behaviors) that are thought to have a direct impact on their overall safety. Each of these elements can be viewed as an input or integrated component in the overall test scenario. The framework was developed in such a way that it could be used for both black-box and white-box testing. Each of the core scenario components can be applied similarly for both black-box and white-box analyses; the differences come in the ability to inject inputs and take output measurements at various levels within the system under test. As part of this analysis, key interfaces where this injection and measurement could take place were identified.

Available literature and reports on current ADS testing activities conducted by both government and industry were reviewed. The review identified three ways that these tests are primarily being conducted:

- Modeling and simulation (M&S)
- Closed-track testing
- Open-road testing

These three techniques offer a multifaceted testing architecture with varying degrees of test control, and varying degrees of fidelity in the test environment. In many cases, two or more of these techniques can be used in parallel or in an iterative fashion to progressively evaluate a complex system such as an ADS.

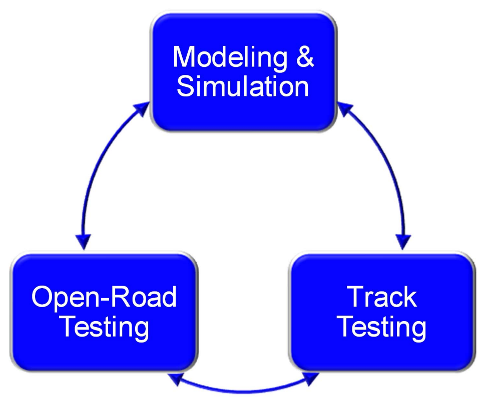

Figure 4. Primary Testing Methods.

Simulation testing provides several advantages:

- Controllability – Simulation affords an unmatched ability to control many aspects of a test.
- Predictability – Simulation is designed to run as specified, so there is little uncertainty as to how the test will run.
- Repeatability – Simulation allows a test to be run many times in the same fashion, with the same inputs and initial conditions.
- Scalability – Simulation allows for generation of a large number and type of scenarios.
- Efficiency – Simulation includes a temporal component, which allows it to be sped up faster than real time so that many tests can be run in a relatively short amount of time.

The following components were identified as collectively making up the core aspects of a common ADS test scenario:

- Tactical maneuver behaviors
- ODD elements
- OEDR capabilities
- Failure mode behaviors

Tactical maneuver behaviors relate to the immediate control-related task(s) the ADS is executing as part of the test (e.g., lane following, lane change, turning). The relevant ODD elements generally define the operating environment in which the ADS is navigating during the test (e.g., roadway type, traffic conditions, or environmental conditions). OEDR capabilities relate directly to the objects and events the ADS encounters during the test (e.g., vehicles, pedestrians, traffic signals). Finally, some tests may include injection or simulation of errors or faults that induce failures at various stages within the ADS's functional architecture.

Test scenarios can be composed of one or more elements of each of these core components, visualized as the individual dimensions of the multidimensional test matrix in Figure 5. Each of these components may be included in a checklist identifying the aspects of each category that are incorporated in a given test.

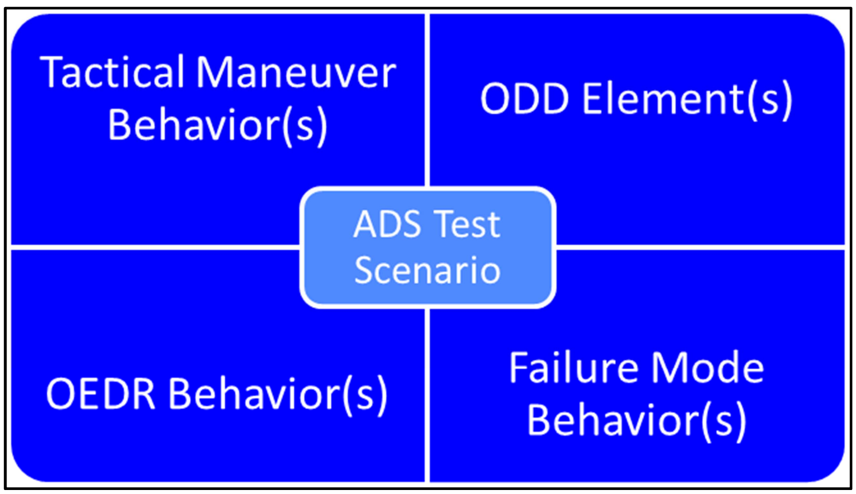

Figure 5. ADS Test Scenario Matrix.

CONCLUSIONS

This paper describes an example of a testing architecture and a scenario-based test framework to support the safe deployment of ADS and evaluate and assess their performance. Efforts focused on the testing of ADS (SAE International L3–L5), where the ADS is fully capable of all aspects of the DDT. To facilitate the identification of the testing architecture and framework, common and relevant operational components for ADS were identified and evaluated, specifically:

- ADS features
- ODD
- OEDR
- FO and FS strategies

The primary contribution of this research is the conceptual development of a test scenario framework that incorporates elements of each of these operational components. The framework uses a checklist-type approach to identify high-level scenario tests by specifying relevant tactical maneuvers, ODD, OEDR, and potential failures.

Each of these components are then further specified to develop a comprehensive set of procedures for a given scenario test. The scenario framework lends itself well to being applied across the three testing techniques identified for the testing architecture (M&S, closed-track testing, and open-road testing), although specific test procedures and implementations will vary, depending on the technique and tools used. This test scenario framework and the sample test procedures developed can provide a launching point to more comprehensive ADS test development and ultimately, test execution. Figure 6 shows a sample ADS test scenario visualization, with the principal elements notionally specified. (In this figure, SV stands for subject vehicle; POV stands for principal other vehicle.)

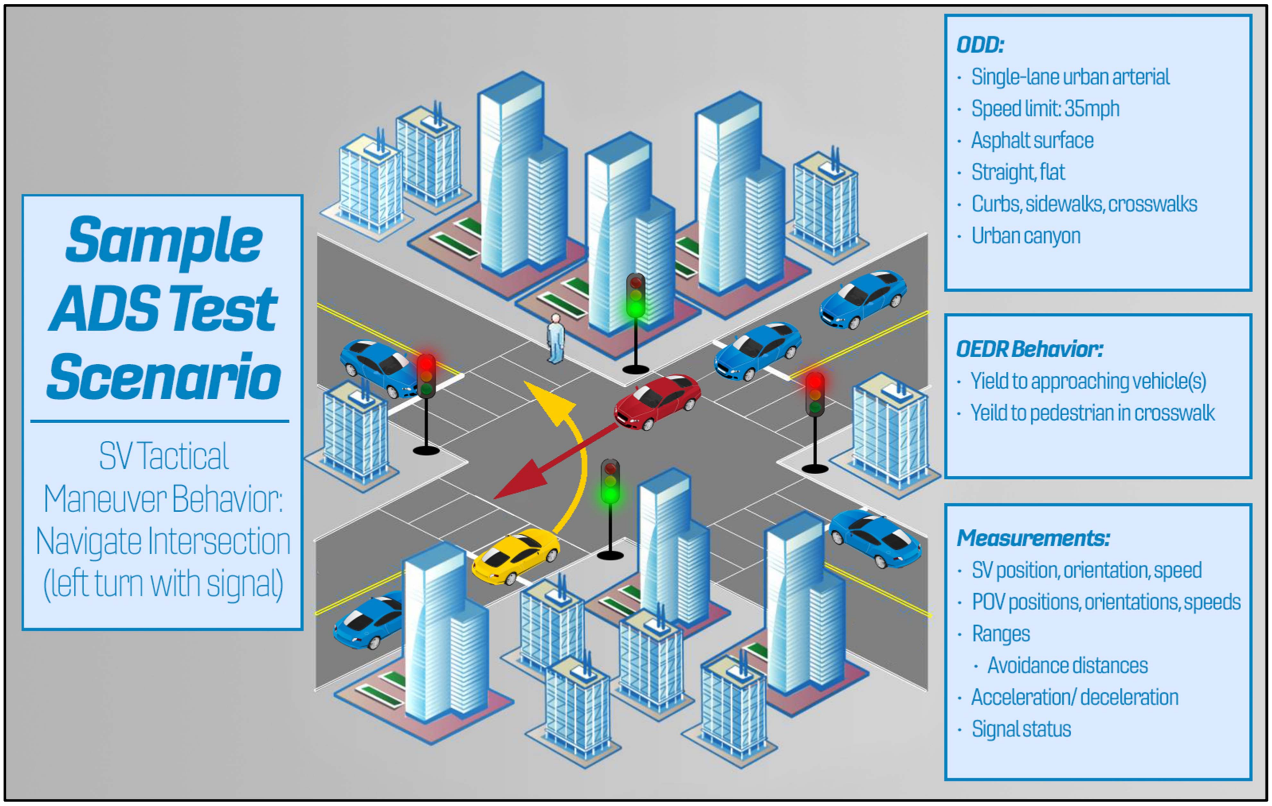

Figure 6. Sample ADS Test Scenario

The expansiveness of conceivable ODD, OEDR, and failure conditions presents a significant challenge to achieving comprehensive testing, even considering the test scenario framework identified during this research and described in this paper. The concept of risk associated with driving scenarios, notionally based on probability and severity of occurrence, has helped focus the analyses of ODD, OEDR, and failure modes to identify an appropriate testing process. A “reasonable worst case” approach may prove sufficient for general safety assessments; however, it is necessary to extend testing beyond the reasonable cases to understand the performance boundaries and limitations of ADS. This paper also identifies M&S capabilities and tools as a potential approach to addressing the expansiveness of these test components, as well as their potential combinations.

REFERENCES

Christensen, A., Cunningham, A., Engelman, J., Green, C., Kawashima, C., Kiger, S., . . . Barickman, F. (2015). Key Considerations in the Development of Driving Automation Systems. *Enhanced Safety Vehicles Conference.* Gothenberg, Sweden: NHTSA .

DARPA. (2008). *Urban Challenge*. Retrieved from DARPA: http://archive.darpa.mil/grandchallenge/

FHWA. (2008). *Managed Lanes: A Primer.*

IIHS. (2016, March). *Automakers agree to standard AEB by 2022*. Retrieved from IIHS HLDI: http://www.iihs.org/iihs/news/desktopnews/u-s-dot-and-iihs-announce-historic-commitment-of-20-automakers-to-make-automatic-emergency-braking-standard-on-new-vehicles

Kyrouz, M. (2017). *Medium*. Retrieved from https://medium.com/smart-cars-a-podcast-about-autonomous-vehicles/industry-comments-to-nhtsas-federal-automated-vehicles-policy-436e7e24911a

NHTSA. (2007). *Pre-Crash Scenario Typology for Crash Avoidance Research.* NHTSA.

NHTSA. (2015). *Human Factors Evaluation of Level 2 and Level 3 Automated Driving Concepts.* NHTSA.

NHTSA. (2016b). *Federal Automated Vehicles Policy.*

NHTSA. (2016c). *Budget Estimates - Fiscal Year 2017.* NHTSA.

NHTSA. (2016d). *Automated Vehicle Research for Enhanced Safety - Final Report.* NHTSA.

NHTSA. (2017b). *Traffic Safety Facts 2015.* NHTSA.

NHTSA. (2017c). *Traffic Safety Facts - 2016 Fatal Motor Vehicle Crashes Overview*. NHTSA.

SAE International. (2018). *J3016: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles.* SAE International.