How Many Operational Design Domains, Objects, and Events?

Philip Koopman, Frank Fratrik

Carnegie Mellon University, Edge Case Research
koopman@cmu.edu, ffratrik@edge-case-research.com

Abstract

A first step toward validating an autonomous vehicle is deciding what aspects of the system need to be validated. This paper lists factors we have found to be relevant in the areas of operational design domain, object and event detection and response, vehicle maneuvers, and fault management. While any such list is unlikely to be complete, our contribution can form a starting point for a publicly available master list of considerations to ensure that autonomous vehicle validation efforts do not contain crucial gaps due to missing known issues.

Introduction

Ensuring that autonomous vehicles will perform adequately in their intended operational environment is a critical part of overall system validation. Traditional software validation includes traceability from requirements to system-level tests. However, the use of machine learning techniques frustrates this approach due to the use of training data rather than a traditional design process. Validation therefore requires at least ensuring that training data and testing covers all relevant operational conditions.

Making this problem tractable in practice is generally accomplished by constraining the operational environment to a subset of all possible situations that could be dealt with by a human driver. That approach to limiting the operational needs of the system is known as adopting an Operational Design Domain (ODD) (NHTSA 2017).

Ensuring that training and testing are complete require at least ensuring that all aspects of the ODD have been addressed either by ensuring safe system operation or by ensuring that the system can recognize and mitigate an excursion beyond the defined ODD. Typical descriptions of an ODD tend to be somewhat simplistic, with NHTSA (2017) listing roadway types, geographic characteristics, speed ranges, weather, and “other domain constraints” as the relevant factors. In our experience with a variety of autonomous vehicle projects we have discovered that the list of “other” considerations can be extensive, and difficult to enumerate without significant experience. Nonetheless, accounting for all of these “other” considerations is essential for ensuring safe real-world operation.

Additional factors to consider in validation include objects and events, generally covered by the term Object and Event Detection and Response (OEDR), which relates to operation within a defined ODD (NHTSA 2017). The term OEDR generally refers to the proper handling of external situations that the vehicle encounters, including perception, planning, and implementation of own-vehicle actions. As with ODD, the NHTSA document provides a short list of possible considerations such as humans manually directing traffic. And again, in practice the list of considerations can become surprisingly large.

Another factor that should be considered is the types of maneuvers the vehicle itself initiates, typically having to do with navigation, such as entering and exiting a limited access roadway, initiating turns, changing lanes, and so on.

Finally, validation should consider responses to and operation with system faults and limitations such as insufficient sensing capability and computational failures. A fault response might include continuing operation with normal capabilities by making use of installed redundancy, reduced capability, or transitioning the system to a safe state. Whichever strategy is chosen, validation must ensure that fault detection and fault responses work properly.

Together these factors result in a four-dimensional validation space with the axes of: {ODD, OEDR, Maneuvers, Fault Management} (Staplin 2018). In general, the cross-product space of all possible factors across all four axes must be addressed.[1] This can be done either by declaring any particular cross-product tuple outside the intended operational space or by ensuring that the system will handle that tuple of factors appropriately. It is tempting to declare some tuples “unlikely” but for a large scale fleet “unlikely” things happen frequently to some vehicle in the fleet. Handling such situations properly or rigorously justifying the improbability of violating a particular operational assumption is an essential part of establishing safety.

[1] The axes selected here conform to existing terminology but have coupling. Desirable future work would be to create an orthogonal set of axes.

Lists of testing scenarios exist, most notably from the PEGASUS project (Lemmer 2017). However, this paper identifies a much richer set of relevant concerns.

The remainder of this paper presents a list of factors we have found relevant to systems for each axis. Handling items on these lists is generally a reasonable expectation of a human driver.

While the list is primarily focused on civilian ground vehicles on public roads, many of the factors are likely to apply to a wide variety of other autonomous systems as well. As extensive as it is, we fully expect that this is only a partial starting point list, and does not include everything that will need to be considered in a road-worthy autonomous car.

ODD Factors

Characterizing the system operational environment should include at least the following:

- Operational terrain, and associated location-dependent characteristics (e.g., slope, camber, curvature, banking, coefficient of friction, road roughness, air density) including immediate vehicle surroundings and projected vehicle path. It is important to note that dramatic changes can occur in relatively short distances.
- Environmental and weather conditions such as surface temperature, air temperature, wind, visibility, precipitation, icing, lighting, glare, electromagnetic interference, clutter, vibration, and other types of sensor noise.
- Operational infrastructure, such as availability and placement of operational surfacing, navigation aids (e.g., beacons, lane markings, augmented signage), traffic management devices (e.g., traffic lights, right of way signage, vehicle running lights), keep-out zones, special road use rules (e.g., time-dependent lane direction changes) and vehicle-to-infrastructure availability.
- Rules of engagement and expectations for interaction with the environment and other aspects of the operational state space, including traffic laws, social norms, and customary signaling and negotiation procedures with other agents (both autonomous and human, including explicit signaling as well as implicit signaling via vehicle motion control).
- Considerations for deployment to multiple regions/countries (e.g., blue stop signs, “right turn keep moving” stop sign modifiers, horizontal vs. vertical traffic signal orientation, side-of-road changes).
- Communication modes, bandwidth, latency, stability, availability, reliability, including both machine-to-machine communications and human interaction.
- Availability and freshness of infrastructure characterization data such as level of mapping detail and identification of temporary deviations from baseline data (e.g., construction zones, traffic jams, temporary traffic rules such as for hurricane evacuation).
- Expected distributions of operational state space elements, including which elements are considered rare but in-scope (e.g. toll booths, police traffic stops), and which are considered outside the region of the state space in which the system is intended to operate.

Special attention should be paid to ODD aspects that are relevant to inherent equipment limitations, such as the minimum illumination required by cameras.

OEDR Factors

System validation should cover at least the following factors, with some factors potentially determined to be out of scope for a particular identified ODD. These can generally be broken down into two sub-categories: objects and events. Specific events might not be applicable if no associated relevant objects are encompassed by the ODD.

OEDR Object Factors

- Ability to detect and identify (e.g. classify) all relevant objects in the environment.
- Processing and thresholding of sensor data to avoid both false positives (e.g., bouncing drink can, steel bridge joint, steel road construction cover plate, roadside sign, dust cloud, falling leaves) and false negatives (e.g., highly publicized partially automated vehicle collisions with stationary vehicles (Orlove 2018))
- Characterizing the likely operational parameters of other road users (e.g., braking capability of leading and following vehicle, or whether another vehicle is behaving erratically enough that there is a likely control fault.)
- Permanent obstacles such as structures, curbs, median dividers, guard rails, trees, bridges, tunnels, berms, ditches, roadside and overhanging signage.
- Temporary obstacles such as transient keep-out zones, spills, floods, water-filled potholes, landslides, washed out bridges, overhanging vegetation, and downed power lines. (For practical purposes, “temporary” might mean obstacles not included on maps, with some vehicle having to be the first vehicle to detect an obstacle for placement even on a dynamic map.)
- People, including cooperative people, uncooperative people, malicious behaviors, and people who are unaware of the operation of the autonomous system.
- At-risk populations which might be unable, incapable, or exempt from following established rules and norms, such as children as well as injured, ability-impaired, or under-the-influence people.
- Other cooperative and uncooperative human-driven and autonomous vehicles.
- Other road users including special purpose vehicles, temporary structures, street dining, street festivals, pa-

rades, motorcades, funeral processions, farm equipment, construction crews, draft animals, farm animals, and endangered species.

- Other non-stationary objects including uncontrolled moving objects, falling objects, wind-blown objects, in-traffic cargo spills, and low-flying aircraft.

OEDR Event Factors

- Determining expected behaviors of other objects, which might involve a probability distribution and is likely to be based on object classification.
- Normal or reasonably expected movements by objects in the environment.
- Unexpected, incorrect, or exceptional movement of other vehicles, obstacles, people, or other objects in the environment.
- Failure to move by other objects which are reasonably expected to move.
- Operator interactions prior to, during, and post autonomy engagement including: supervising driver alertness monitoring, informing occupants, interaction with local or remote operator locations, mode selection and enablement, operator takeover, operator cancellation or redirect, operator status feedback, operator intervention latency, single operator supervision of multiple systems (multi-tasking), operator handoff, loss of operator ability to interact with vehicle.
- Human interactions including: human commands (civilians performing traffic direction, police pull-over, passenger distress), normal human interactions (pedestrian crossing, passenger entry/egress), common human rule-breaking (crossing mid-block when far from an intersection, speeding, rubbernecking, use of parking chairs, distracted walking), abnormal human interactions (defiant jaywalking, attacks on vehicle, attempted carjacking), and humans who are not able to follow rules (children, impaired adults).
- Non-human interactions including: animal interaction (flocks/herds, pets, dangerous wildlife, protected wildlife) and delivery robots.

Maneuvers

While vehicle operations are often discussed in terms of maneuvers, this category in practice must expand to include other aspects of operation that go beyond controlling vehicle motion itself. Relevant aspects include:

- Operational actions, maneuvers, direction of travel, path planning, goal setting, and goal seeking behaviors. This generally includes various vehicle geometries, and various driving behaviors such as turns, lane changes, exits, entrances, parking, and so on.
- Mission length and mission profile (e.g., whether a secondary safing mission is used as a response to a malfunction, unoccupied operations).
- Operational modes and safe transition between modes, including: power on/self-test, autonomous operation, human-directed operation, safe state operation, maintenance (fueling, repair, car wash, consumable replacement, cleaning, calibration), transportation, fault response, post-fault response (e.g., to ensure emergency responder safety after a mishap), fault diagnosis, update validation, and conformance testing.
- Change in ownership and change in operational profile (e.g., relocation, redeployment, overhaul, upgrade).

Fault Management

While traditional functional safety approaches include many aspects of fault management, they do not necessarily deal with requirements gaps and ensuring safety when the system encounters an environmental exception or other situation for which it was not designed. Moreover, with the removal of a human operator autonomy can be burdened with detecting, diagnosing, and mitigating faults that would otherwise be handled by a human driver.

We identify the sub-categories of system limitations, system faults, and fault responses.

System Limitations

- Current capabilities of sensors and actuators, which can depend upon the operational state space.
- Detecting and handling a vehicle excursion outside the operational state space for which it was validated, including all aspects of {ODD, OEDR, Maneuver, Fault} tuples.
- Desired availability despite fault states, including any graceful degradation plan, and any limits placed upon the degraded operational state space.
- Capability variation based on payload characteristics (e.g. passenger vehicle overloaded with cargo, uneven weight distribution, truck loaded with gravel, tanker half filled with liquid) and autonomous payload modification (e.g. trailer connect/disconnect).
- Capability variation based on functional modes (e.g. pivot vs. Ackerman vs. crab steering, rear wheel steering, ABS or 4WD engaged/disengaged).
- Capability variation based on ad-hoc teaming (e.g. V2V, V2I) and planned teaming (e.g. leader-follower or platooning vehicle pairing).
- Incompleteness, incorrectness, corruption or unavailability of external information (V2V, V2I).

System Faults

- Perception failure, including transient and permanent faults in classification and pose of objects.
- Planning failures, including those leading to collision, unsafe trajectories (e.g., rollover risk), and dangerous paths (e.g., roadway departure).
- Vehicle equipment operational faults (e.g., blown tire, engine stall, brake failure, steering failure, lighting system failure, transmission failure, uncommanded engine power, autonomy equipment failure, electrical system failure, vehicle diagnostic trouble codes).
- Vehicle equipment maintenance faults (e.g., improper tire pressure, bald tires, misaligned wheels, empty sensor cleaning fluid reservoir, depleted fuel/battery).
- Operational degradation of sensors and actuators including temporary (e.g., accumulation of mud, dirt, dust, heat, water, ice, salt spray, smashed insects) and permanent (e.g., manufacturing imperfections, scratches, scouring, aging, wear-out, blockage, impact damage).
- Equipment damage including detecting and mitigating catastrophic loss (e.g., vehicle collisions, lighting strikes, roadway departure), minor losses (e.g., sensor knocked off, actuator failures), and temporary losses (e.g., misalignment due to bent support bracket, loss of calibration).
- Incorrect, missing, stale, and inaccurate map data.
- Training data incompleteness, incorrectness, known bias, or unknown bias.

Fault Responses

Some of the faults and limitations fall within the purview of safety standards that apply to non-autonomous functions. However, a unified system-level view of fault detection and mitigation can be useful to ensure that no faults are left unaddressed. More importantly, to the degree that credit has been taken for a human driver participating in fault mitigation by safety standards, that places fault mitigation obligations upon the autonomy.

- How the system behaves when encountering an exceptional operational state space, experiencing a fault, or reaching a system limitation.
- Diagnostic gaps (e.g., latent faults, undetected faults, undetected faulty redundancy).
- How the system re-integrates failed components, including recovery from transient faults and recovery from repaired permanent faults during operation and/or after maintenance.
- Response and policies for prioritizing or otherwise determining actions in inherently risky or certain-loss situations.
- Withstanding an attack (system security, compromised infrastructure, compromised other vehicles), and deterring inappropriate use (e.g., malicious commands, inappropriately dangerous cargo, dangerous passenger behavior).
- How the system is updated to correct functional defects, security defects, safety defects, and addition of new or improved capabilities.

Conclusions

This is clearly a long list of things to consider. And yet, it is almost certainly incomplete. Nonetheless, it is a starting point for further discussion about what types of issues should be included in a minimum performance validation approach. An essential next step will be finding ways to manage the combinatorial complexity of validation without missing emergent effects that cause some combinations of factors to cause unexpected and dangerous results even if individual underlying aspects of that combination have all been individually addressed.

It is important to realize that even getting every single relevant item on these lists right is insufficient to establish safety. While these lists hopefully cover a significant fraction of safety relevant concerns as well as basic functionality concerns, each vehicle should be evaluated for safety in the context of its architecture, design, implementation, and intended use. The good news is that many vehicles won't have to deal with every element on these lists if they are able to adopt limited ODDs.

A primary purpose of the lists given is to ensure that there are no surprises due to omitting factors that are known to be relevant but have somehow been overlooked. It is hoped that over time the autonomous vehicle industry will begin to share information such as this to avoid unnecessary loss events due to overlooking an issue that was already known to be a problem.

References

Lemmer, K., PEGASUS: Effectively ensuring automated driving, April 6, 2017. https://www.pegasusprojekt.de/files/tmpl/pdf/PEGASUS_VDA_techn.congress_EN.pdf accessed November 3, 2018.

NHTSA, 2017, *Automated Driving Systems: a vision for safety*, US Dept. of Transportation, DOT HS 812 442, Sept. 2017.

Orlove, R., "This test shows why Tesla Autopilot crashes keep happening," Jalopnik, June 13, 2018, https://jalopnik.com/this-test-shows-why-tesla-autopilot-crashes-keep-happen-1826810902 accessed November 3, 2018.

Staplin, L., Mastromatto, T., Lococo, K. H., Kenneth W. Gish, K. W., & Brooks, J. O. (2018, September). The effects of medical conditions on driving performance (Report No. DOT HS 812 623). National Highway Traffic Safety Administration