Identifying Effective *STPA Control Structures to Characterize SOTIF Areas 1,2,3, and 4 in Automated Vehicles.*

Xuezhu, Yang

Zhongju, Di

China FAW Corporation Limited

China

Juan Pimentel

Rolf Johansson

Greg Gruska

OMNEX (US) Co. Ltd

US

Ruoyu, Xu

Fu, Xu

OMNEX (Shanghai) Co. Ltd

China

Paper Number 23-0019

ABSTRACT

Developing and automated driving system (ADS) for an automated vehicle with a sufficient level of safety has turned out to be a much more difficult problem than anticipated by the industry. The challenges are multiple, for example the existence of a very large number of critical scenarios that would require testing vehicles for billions of miles to guarantee safety. In this paper we propose using System Theoretic Process Analysis, STPA, to characterize SOTIF areas 1, 2, 3, and 4 for SAE automation levels 3 and 4. A key challenge of STPA is the identification of an appropriate dynamic control structure that is efficient for the purpose at hand. We propose a control structure built around the decision hierarchy of strategical, tactical, and operational decisions, used to structure an ADS including its relations to the user, the environment, and all other traffic actors. More specifically, we show how an analysis based on this control structures at the strategic, tactical, and operational levels can be used to identify safe and unsafe control actions (UCAs) in known scenarios

INTRODUCTION

It is widely accepted that an automated driving system (ADS) operates in a very complex environment and this one of the reasons why commercial deployments of safe autonomous vehicles has turned out to be much more difficult than expected. Currently there is a international standard ISO 21448, also called the Safety of the Intended Functionality (SOTIF), that addresses the safety of the intended functions of ADSs [1]. An ADS operates in a wide variety of scenarios which must be carefully characterized for a variety of purposes, e.g., conceptual design, implementation, testing, verification, validation, etc. [2]. Depending upon whether scenarios are known or unknown and whether such scenarios cause hazardous behaviour or not, SOTIF classifies scenarios into four areas called Area 1 through Area 4. Safety analysis of each of the SOTIF areas helps determine whether an ADS design has an acceptable level of risk. Recently, there has been much work on system theoretic process analysis (STPA) as a safety analysis method in various contexts, including SOTIF [3-7]. One important step of the STPA method involves the development of control structures [8]. In this paper we identify an effective STPA control structure that would prove useful for the

SOTIF safety analysis of Areas 1, 2, 3, and 4 in the context of an ADS project involving a highway pilot (HWP) feature of an SAE level 3 automation system as defined in the standard SAE J3016 [9]. This paper focuses on the STPA control structure while a future paper will address the safety analysis of SOTIF areas 1,2,3, and 4 based on the proposed control structure.

The contributions of this paper are as follows: a) the development of a detailed STPA control structure that is hierarchical in nature taking into account the three levels specified by the SAE J3016 standard namely the strategical, tactical, and operational control levels, b) a set of unsafe control actions which were identified for the control structure, and c) a discussion on the effectiveness of the STPA Control Structure to characterize SOTIF areas 1,2,3, and 4.

A. The SAE J3016 Automation and Hierarchical Control Levels

The SAE J3016 document [9] defines 6 levels of automation that are possible in terms of the performance of lateral and longitudinal automation, the monitoring of the driving environment, the fallback when automation fails (also called the DDT (dynamic driving task) fallback) and the scope of the ODD (operational design domain):

- Level 0 (No automation). As the name implies, at this level there is actually no automation.
- Level 1 (Driver assistance). At this level there is either lateral or longitudinal automation (but not both). A driver monitors the driving environment and must be ready to take over when automation fails. The ODD is limited.
- Level 2 (Partial automation). At this level there is both lateral and longitudinal automation. A fallback driver monitors the driving environment and must be ready to take over when automation fails. The ODD is limited.
- Level 3 (Conditional automation). At this level, the ADS provides lateral and longitudinal automation and in addition it monitors the driving environment. A user must be ready as a fallback when automation fails. The ODD is limited.
- Level 4 (High automation). At this level, the ADS provides lateral and longitudinal automation, it monitors the driving environment, and acts as a fallback when automation fails. The ODD is limited.
- Level 5 (Full automation). At this level, the ADS provides lateral and longitudinal automation, it monitors the driving environment, and acts as a fallback when automation fails. The ODD is unlimited.

While discussing the DDT, the J3016 document makes reference to a three-level control schematic shown in Fig. 1 consisting of three levels: strategic, tactical, and operational [9].

According to J3061, the DDT includes the following sub-tasks:

1. Lateral vehicle motion control via steering (operational).
2. Longitudinal vehicle motion control via acceleration and deceleration (operational).
3. Monitoring the driving environment via object and event detection, recognition, classification, and response preparation (operational and tactical).
4. Object and event response execution (operational and tactical).
5. Maneuver planning (tactical).
6. Enhancing conspicuity via lighting, sounding the horn, signaling, gesturing, etc. (tactical).

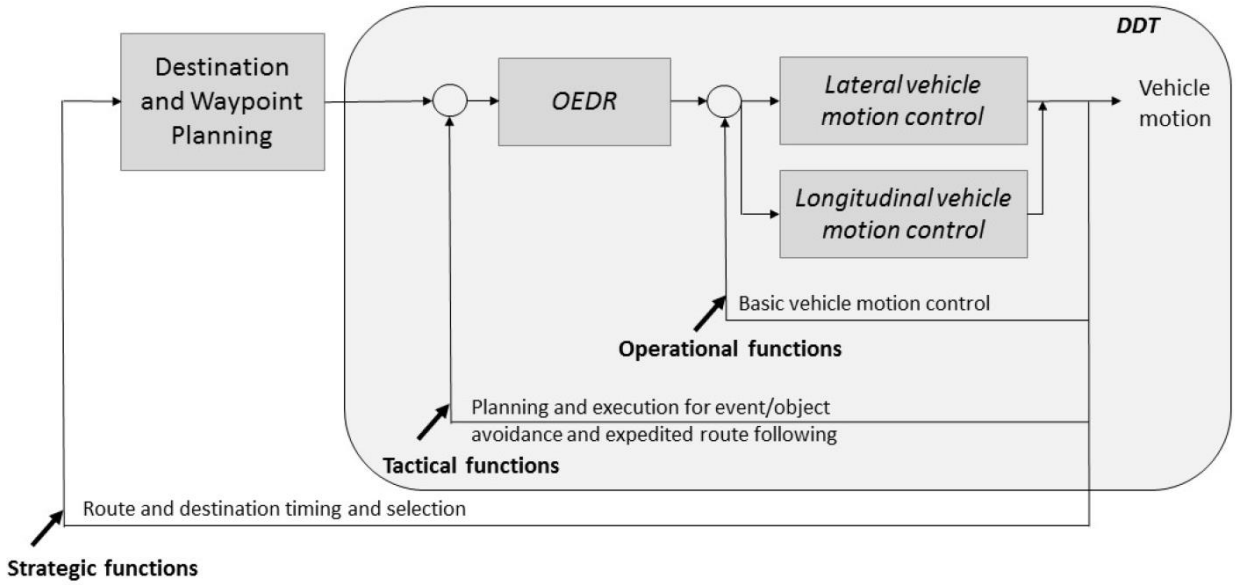

Figure 1. Schematic view of the driving tasks of a DDT per SAE J3016.

To enable a DDT with a sufficient level of safety and performance, appropriate models are necessary that include:

1. Models of the driving environment
2. Models for DDT fallback
3. Models for object and event detection, recognition, classification, and response preparation
4. Models for object and event response execution
5. Models for maneuver planning

In this paper, we provide a specific model of the driving environment in the context of a J3016 based hierarchical STPA control structure. We use the control structure together with the driver environment model (DEM) to perform hazard analysis and to identify SOTIF areas 1 and 2 and to provide an in-depth analysis of areas 3 and 4. Although not explicitly addressed, this paper provides discussions on a few aspects of the other models 2 through 5 in the above list.

B. SOTIF Areas

The SOTIF standard DIS ISO 21448 [1] classify scenarios into four categories as depicted in the left side of Fig. 2 where only Areas 1 and 2 are known and only Areas 2 and 3 are hazardous. According to this draft international standard, the areas are conceptual abstractions representing a goal of the SOTIF process, which is to:

- Perform a risk acceptance evaluation of Area 2 based on the analysis of the intended functionality.
- Reduce the probability of known scenarios causing hazardous behavior, in Area 2, to an acceptable level of risk.
- Reduce the probability of the unknown scenarios causing potentially hazardous behavior, in Area 3, to an acceptable level of risk.

The presence of unreasonable risk might be evident at the beginning of the development, visualized by too large Areas 2 and 3. The ultimate goal of the SOTIF activities is to evaluate the potentially hazardous behavior present in Areas 2 and 3 and to provide an argument that these areas are minimal, i.e., at or below acceptance criteria, and therefore the residual risk caused by these scenarios is sufficiently low.

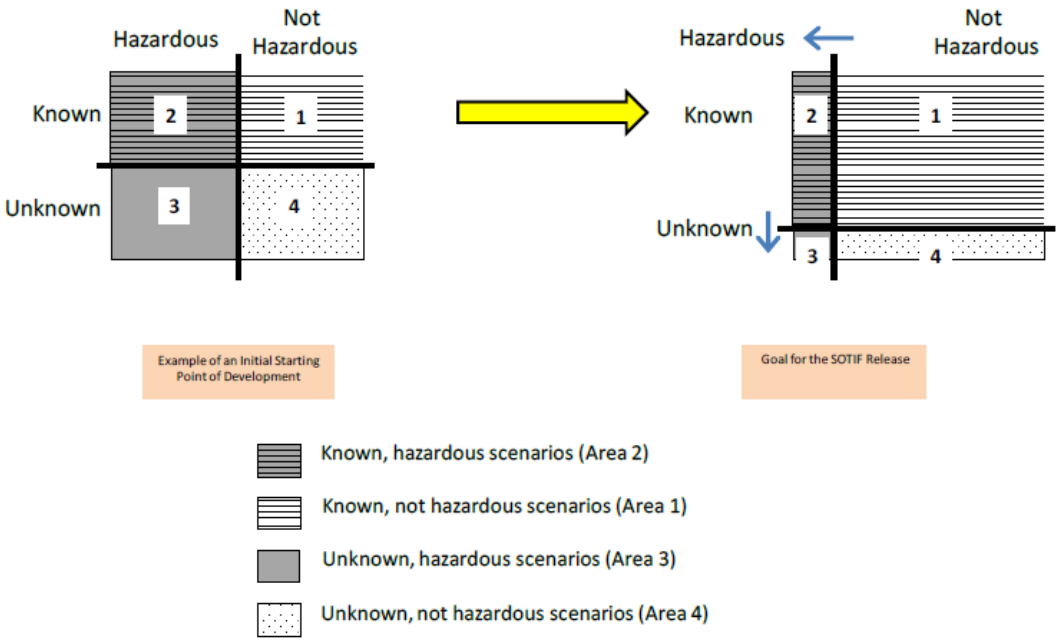

Figure 2. SOTIF areas and its evolution resulting from the ISO 21448 activities.

C. The STPA process

System theoretic process analysis (STPA) is a safety analysis framework based in control system theory with applications in many industrial areas including automotive. The STPA process consists of 4 steps [8]:

- Step 1. Define the purpose of the analysis
- Step 2. Model the control structure
- Step 3. Identify unsafe control actions (UCAs)
- Step 4. Identify loss scenarios

In step 1, the system under design is characterized as much as possible. Being a framework based on control systems theory, specific control structures are modeled in step 2. Based on the specific control structures the STPA process models hazards as unsafe control actions which are outcomes of the various controllers in the control structure. In step 4, the fundamental question as to how is it possible for the controllers to generate UCAs is answered through the identification of causal factors in the context of specific scenarios. This paper only focuses on STPA steps 2 and 3.

The STPA Hierarchical Control Structure

When using methods such as STPA for hazard analysis in an ADS development process, it is important to introduce structure as a means to deal with such complex environment. This structure needs to be detailed enough to enable the capture of what is required by SOTIF; examples of such detail include triggering conditions, insufficiency of specification, performance limitations, and reasonably foreseeable misuse. The structure discussed in this paper is the STPA control structure which is an important step of the STPA process.

In addition to structuring the control, an STPA control structure can be also used to structure additional entities such as:

- Hierarchical control levels
- Information and control flows
- Feedback control loops
- Loss scenarios
- Functional insufficiencies
- Misuse
- Refined hazards

- Safety requirements
- Test scenarios [10]

Furthermore, we need to reduce the scope of the complex environment to only cover what is strictly necessary for the feature in question, e.g., highway pilot. This reduction in scope and complexity will limit, among other things, the number of scenarios and this will enable to perform a more complete safety analysis of the ADS and in turn to characterize SOTIF areas 1, 2, 3, and 4 more precisely.

The implication for an ADS is that on the one hand, the very complex environment including the driver needs to be captured in the control structure, such that the interaction with other traffic actors can be covered. On the other hand, the STPA control structure should also capture the essence of what it takes to act autonomously in the context of the overall driving environment including the main actors such as an operator, a fallback ready user, and other vehicles.

The proposed STPA control structure (i.e., step 2 of the STPA process) is shown in Fig. 3. As depicted in this Fig., in addition to including the hierarchical levels of J3016, the developed STPA Control Structure includes a detailed driving environment model (DEM) to be described in a future paper. Such model will enable a detailed safety analysis with the goal of an accurate characterization of known and unknown scenarios (SOTIF areas 1 to 4). Furthermore, the STPA control structure includes information related to the sensors and perception system, the control algorithms, the actuators, and the controlled process. For clarity, only a few examples of control actions and feedback are depicted in Fig. 3. Note that the controlled process includes the ego vehicle and the physical and driving environment.

Information from the physical and driving environment is captured by appropriate sensors external to the vehicle as part of a perception system composed of cameras, radars, Lidars, IMU (inertial measurement unit) and ultrasonic sensors. The output of these sensors are inputs to a perception fusion system that congregates information about objects in an integrated object model. The HMI block constitutes an interface between the ADS and the operator which in the case of a level 3 ADS is the fallback ready user. There is also a localization and mapping block that is in charge of determining the exact location of the vehicle in real world coordinates by making use of high definition (HD) maps and appropriate GPS, GNSS, and related blocks that is assumed to be available to the ADS. When enabled by the tactical controller a vehicle trajectory is generated and acts as the reference path for the operational controller to control the detailed movement of the ego vehicle in the driving environment. The operational controller ensures that the commanded trajectory is followed as close as possible by using longitudinal and lateral controls. We assume that the ADS is built on top of an already functioning vehicle, thus much of the operational controller is already existing and available.

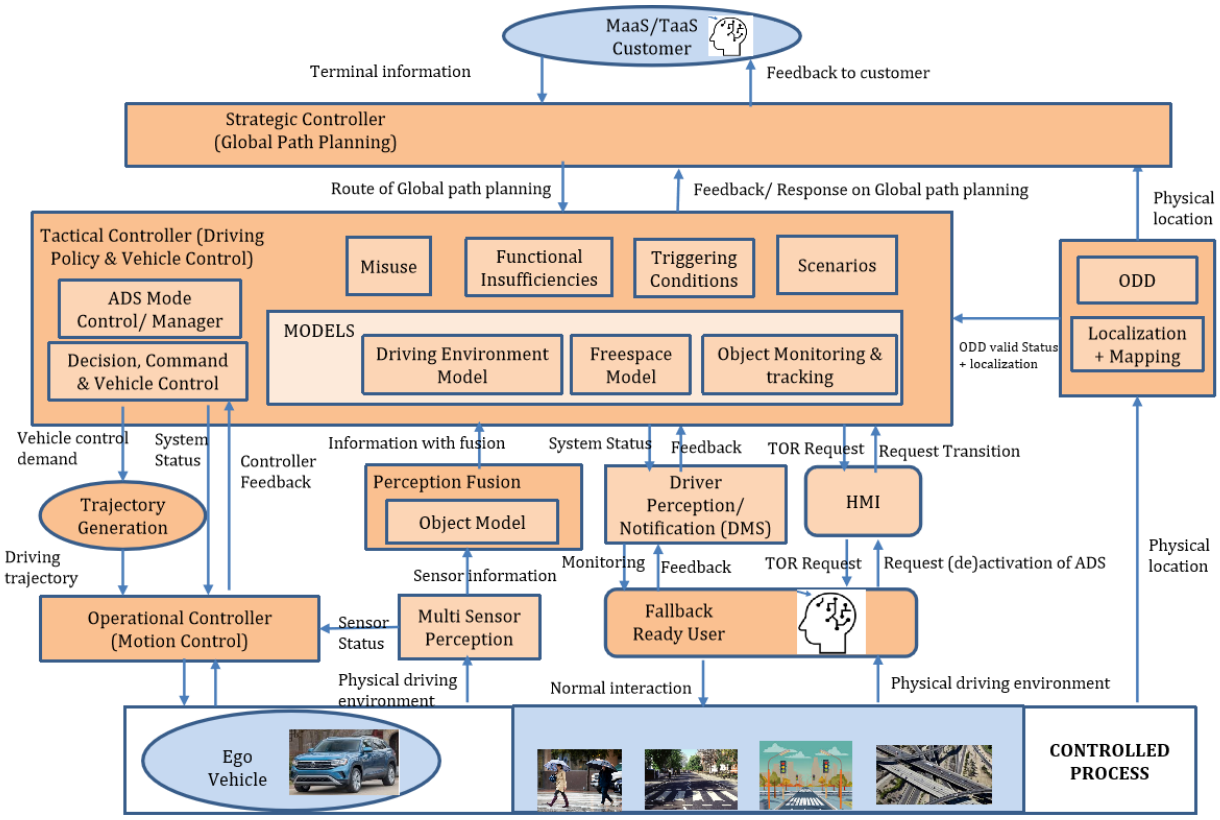

Figure 3. Proposed STPA control structure for a level 3 automated driving system (ADS). The up arrows (↑) indicate feedback, the down arrows (↓) indicate control actions, and the side arrows (←) indicated other type of information.

Whereas other authors consider the physical and driving environment as an input to the controllers [10], we put them as part of the controlled process which includes inputs and outputs. There are several reasons for this decision. First, through control actions, the various controllers effect changes on the environment, e.g., a behavioral change. Second, through the ego vehicle it is possible to effect change on the driving environment such as enhancing conspicuity via lighting, sounding the horn, signaling, gesturing, etc. In the future, this degree of control of the environment will increase with the introduction of wireless communications in the context of V2V (vehicle to vehicle), V2X (vehicle to infrastructure), and V2P (vehicle to pedestrian).

The operational design domain (ODD) is an important element of the STPA control structure. In this paper, we use the ODD structure proposed by the NHTSA consisting of the following elements [11]: (1) Physical infrastructure, (2) operational constraints, (3) objects, (4) connectivity, (5) environmental conditions, and (6) zones. The physical infrastructure includes such things as roadway types, roadway surfaces, roadway edges, and roadway geometry. Operational constraints include things like speed limit and traffic conditions. Likewise, Objects include things like signage, roadway users, non-roadway users, obstacles, objects, etc. Connectivity includes vehicles, traffic density information, remote fleet management systems, infrastructure sensors and communications. Environmental conditions include weather, weather induced road conditions, particulate matter, and illumination and zones include geo-fencing, traffic management zones, school construction zones, regions, states, and interference zones.

The ODD structure described above is generic to be used by all possible vehicle features. In the specific driver environment model described below, only a few elements of the overall ODD structure is used in order to have a simple and specific DEM that is effective for safety analysis and design of a specific level 3 ADS feature, that of an HWP.

A. Strategic Controller

The main task of the strategical decision-level controller is to define the goal of the trip. This includes a negotiation with the mobility-as-a-server (MaaS) / Transport-as-a-Service (TaaS) user, but it also includes the alternatives of never start, and of interrupting a trip changing its strategical decision to a minimal risk condition, MRC. This means that the MaaS/TaaS user might come with preferred trip destinations, but it is the ADS strategical controller that makes the decision of what is a safe strategical control action on this level, i.e. formulates the safe control action for the tactical decision level controller to execute. It is important that the strategical-level controller decision is with the ADS, to be able to avoid any unsafe control action. It is fundamental to formulate constraints on the strategical decision level controller, not to accept any strategically decision that cannot be guaranteed to be able to reach safely. As this is a control-loop responding to feed-back, the strategical decision needs to be reassured constantly, which for example may lead to either suggesting the user to take back the control or to change the ADS strategical decision to an MRC. In the full paper we show the full set of unsafe control actions for the strategical controller, and how these can be connected to a limited set of loss scenarios.

B. Tactical Controller

In terms of safety assurance of the HWP feature at an SAE level 3, the tactical controller is the most critical controller in the hierarchy. Although there are hazards at the strategic and operational controller hierarchical levels, the nature of such hazards is different from those of the tactical controller. More specifically, hazards at the strategic controller are out of scope of a DDT and thus of this paper. There are basically two categories of hazards at the operational level; the first category are hazards due to malfunctions and this is addressed by functional safety (i.e., ISO 26262); the second category of hazards are SOTIF in nature and should be addressed in a SOTIF analysis but is out of scope of this paper because the focus is on the tactical controller and an underlying model of the driving environment.

The tactical controller is depicted in Fig. 4 include the following main block categories:

1. Feedback and Inputs
2. Models
3. Decision making (e.g., control algorithms, driving laws)
4. Controller Actions / Execution

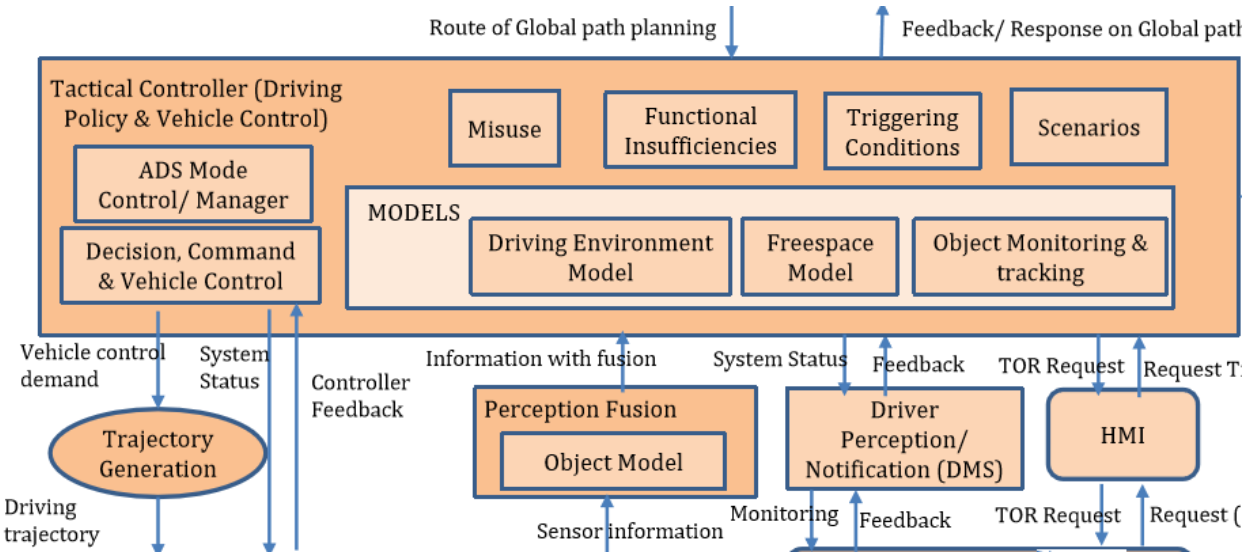

Figure 4. Constituent components of a tactical controller for a level 3 automated driving system (ADS).

C. Operational Controller

On the operational decision level, we find the traditional controller model, that for an ADS will be to execute a given trajectory. The controller constraints to avoid unsafe control actions, will include what it takes to be sufficiently close to the given trajectory. Whether this trajectory is safe or not, is not a question for the operational level controller where the only task is to execute a trajectory that is deemed safe by the controller on higher level. As for all the controllers, this is a continuous process where new decisions are given and new feed-back from the controlled process appears all

the time. The important thing is that the controller constraints are identified such that the feed-back loops of the controllers can guarantee the ADS to stay safe. By limiting the task for the operational controller to what is complementary to the two other controllers, it is possible to limit the set of loss scenarios, and thus to formulate a feasible verification strategy avoiding the 'billion miles problem'.

Effectiveness of the STPA Control Structure to characterize SOTIF areas 1,2,3, and 4

Scenarios are complex entities which are crucial to SOTIF analysis and there has been much discussion and interest in characterizing them. One approach to characterize scenarios is to develop a set of scenarios based on the features and use cases and utilize these scenarios as input for the safety analysis. One issue with this approach is that there is potentially a very large number of scenarios to consider. A related issue is the question as to what extent all scenarios are relevant for safety analysis, particularly the ones that are safety critical). In this context, the scenarios that are relevant but still not part of the analysis, are what constitutes SOTIF area 3. Another approach to characterize scenarios is to consider them not as inputs to the safety analysis but rather as a result of a complex interplay between the environment and the ego vehicle. In other words, scenarios can be viewed as outcomes of dynamic interactions of the STPA control structure and the controlled process which includes the ego vehicle and the physical and driving environment. The question about completeness of the scenarios, is a question about how to reach completeness for the purpose of analysis. If one is not using the scenarios as independent inputs, but as a result of a complex interplay between the environment and the ego vehicle, the problem of reaching completeness is no longer theoretically impossible. In this context, reaching completeness means not leaving anything out of the analysis. In our work we assume the latter approach described above. Regardless, the STPA control structure plays a crucial role in the safety analysis of an ADS in the context of SOTIF areas 1 to 4. Details will be provided in a future paper by the authors.

Results

In this section we provide some preliminary results of using the proposed STPA control structure depicted in Figs. 3 and 4 in an actual design of an automated vehicle manufactured by a big OEM. We present results for a HWP (highway pilot) feature at an SAE automation level 3. Through the efficient STPA control structure proposed in this paper we have identified 13 system level hazards and 24 Control flows (considering only the strategic and operational controllers).

A partial list of unsafe control actions includes:

1. HWP ADS does not provide (or providing the incorrect) the longitudinal or lateral control action during cruise control
2. HWP ADS provide longitudinal or lateral control during cruise control when it is not requested
3. HWP ADS does not provide the longitudinal or lateral control action during unplanned situations while driving
4. HWP ADS provide the unnecessary longitudinal or lateral control actions during unplanned situation while driving
5. HWP ADS provides too much or too little longitudinal or lateral control action during unplanned situations while driving
6. HWP ADS provides too early or too late or out of order longitudinal or lateral control action during unplanned situations while driving
7. HWP ADS does not provide the longitudinal or lateral control action while avoiding objects on the highway
8. HWP ADS provide unnecessary longitudinal or lateral control actions while no objects to avoid
9. HWP ADS does not provide the longitudinal or lateral control action while maintaining commanded trajectory
10. HWP ADS provide unnecessary longitudinal or lateral control while maintaining commanded trajectory

Summary and Conclusions

We have detailed a hierarchical STPA control structure having three hierarchical levels to generate scenarios suitable for safety analysis of SOTIF areas 1, 2, 3, and 4. Scenarios are complex entities that include features and events occurring over the operating lifetime of the vehicle and also include circumstances in which the hazard can lead to harm. The hierarchical control structure is based on the strategic, tactical, and operational functions defined in SAE J3016. Decomposing the responsibilities of achieving the DDT between the tactical and operational controllers

together with the analysis of the high-level system hazards enable the identification of unsafe control actions which are correlated to SOTIF requirements.

REFERENCES

[1] ISO 21448 - Road vehicles— Safety of the Intended Functionality.

[2] Till Menzel, Gerrit Bagschik, and Markus Maurer. Scenarios for development, test and validation of automated vehicles. In 2018 IEEE Intelligent Vehicles Symposium (IV), pages 1821–1827. IEEE, 2018.

[3] A. Abdulkhaleqa, et al., 'A Systematic Approach Based on STPA for Developing a Dependable Architecture for Fully Automated Driving Vehicles', 4th European STAMP Workshop 2016.

[4] Abdulkhaleq, Asim & Wagner, Stefan & Lammering, Daniel & Boehmert, Hagen & Blueher, Pierre. (2017). Using STPA in Compliance with ISO 26262 for Developing a Safe Architecture for Fully Automated Vehicles. arXiv:1703.03657v1 [cs.SE] 10 Mar 2017.

[5] M. Chaal et al., 'A framework to model the STPA hierarchical control structure of an autonomous ship', in Safety Science, Volume 132, December 2020.

[6] Zhang, S.; Tang, T.; Liu, J. A, Hazard Analysis Approach for the SOTIF in Intelligent Railway Driving Assistance Systems Using STPA and Complex Network. Appl. Sci. 2021, 11, 7714. https://doi.org/10.3390/app11167714

[7] S. M. Sulaman, et al, 'Hazard Analysis of Collision Avoidance System using STPA', in Proceedings Information Systems for Crisis Response And Management (ISCRAM) , 2014.

[8] N.G. Leveson, J.P. Thomas, 'STPA Handbook, MIT, March 2018, http://psas.scripts.mit.edu/home/materials/

[9] SAE J3016:APR2021, Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles, April 2021.

[10] S. Khastgir et al. 'Systems Approach to Creating Test Scenarios for Automated Driving Systems', in Reliability Engineering & System Safety, volume 215, November 2021.

[11] Eric Thorn, Shawn Kimmel, Michelle Chaka (2018, September). *A Framework for Automated Driving System Testable Cases and Scenarios* (Report No. DOT HS 812 623). Washington, DC: National Highway Traffic Safety Administration.