1.2 – Identifying hazardous system behaviour

Practical guidance – automotive

Author: Dr Richard Hawkins, Assuring Autonomy International Programme

This assurance objective requires the identification of behaviour at the vehicle-level that may (in some situations) result in a hazard. For autonomous driving this will often be referred to as a hazardous event [1], which is a combination of a particular operational situation with a hazard. This therefore firstly requires a definition of the relevant operational situations (Section 1.1.3 of the BoK provides a discussion of how operational scenarios may be defined for autonomous driving). The hazards that may arise in these situations must then be identified by looking at the deviations in the vehicle behaviour that may occur. The focus of deviations is often on component malfunction, however for autonomous driving deviations may result from a wide range of causes including:

- systematic errors in the functional specification and design
- interaction failures between vehicle and its environment or driver
- operation outside of the defined ODD (see section 1.1.2)

These sources of deviation must also be considered as part of the analysis.

ISO 26262 [1] provides some limited guidance on performing vehicle level hazard analysis, suggesting techniques such as brainstorming, checklists, FMEA and field studies. None of these techniques have been found to be particularly effective for autonomous driving systems mainly due to the number and variety of operational situations that must be considered. Instead, the two most commonly adopted techniques for hazard analysis for autonomous driving are HAZOP and STPA.

HAZOP

HAZOP [2] is a technique developed for analysing system hazards by considering potential deviations in the system behaviour. The deviations are identified by combining items of the system with a set of specified guidewords. Originally HAZOP was developed for analysing chemical plants, where the items under consideration were flows of materials through the system. HAZOP has since been extended and applied to a wide range of other domains including automotive, where functional items are often the focus of the analysis. The success of applying HAZOP hinges on the interpretation of guidewords to identify hazardous behaviour. The standard HAZOP approach provides a large number of guidewords that may be interpreted and tailored to the system that is being analysed. There is currently little guidance on how to modify or interpret guidewords effectively for autonomous driving, or how to explicitly consider the impact of the scenarios and situations that the vehicle may encounter on the safety of the outcome.

One approach for using HAZOP in autonomous driving is described in [3]. This defines keywords for different categories of vehicle activity, either perception, planning or action. Keywords proposed for the different categories are shown below:

- Perception: no, non-existent, erroneous, too large, too small
- Planning: not relevant, relevant {parameter e.g. object} not, conflicting, physically not possible
- Action: absent, wrong, unattended, too large, too small

This set of guidewords should be taken as guidance only, different or additional guidewords should be created where necessary.

These guidewords are combined with scene descriptions in order to identify hazardous events. An example output of this process (taken from [3]) is shown in Figure 1 for the function of "select relevant object". The hazardous event is identified as a malfunction within the defined scenario.

Mode	Follow mode
Function	Select relevant object
Malfunction	Relevant object not considered
Road infrastructure	Solid line (left) and turf (right)
Object constellation	Vulnerable object
Curvature, width and weather	Valid
Traffic constellation	Moving traffic with no limitation
Driving state	Driving at 10 km/h

Figure 1 - Example hazardous event identified using HAZOP (taken from [3])

This approach requires a definition of the scenarios that are to be considered. Determining a suitable level of abstraction of scenario definition to support the analysis is an open challenge (see section 1.1.3).

HAZOP analysis can be applied at different levels of abstraction and stages of the lifecycle. A detailed example is provided in [4] of the application of HAZOP to an automated lane centering (ALC) system. The following 7 guidewords were are applied to each of the 24 functions of the components of the ALC system:

- Loss of function
- More than intended
- Less than intended
- Intermittent
- Incorrect direction
- Not requested
- Locked function

This provides a very detailed analysis at a component level and generates a large set of hazardous malfunctions (113 for the ALC). This application of HAZOP is less focussed on vehicle-level behaviour but rather on component malfunction.

STPA

STPA (Systems-Theoretic Processes Analysis) [5] is a technique that uses a systems-theoretic accident model to identify causal factors and unsafe scenarios that may result in a hazardous outcome. The main steps of STPA are:

1. Establish the fundamentals of the analysis (e.g. system-level accidents and the associated hazards) and draw the control structure diagram of the system.
2. Use the control structure diagram to identify the potentially unsafe control actions.
3. Determine how each potentially unsafe control action could occur by identifying the process model and its variables for each controller and analysing each path in the control structure diagram.

STPA is suited to the identification of hazardous behaviour for autonomous driving due to its focus on the dynamic control of a system and on causes of hazards in the absence of failure. [6] provides an example of how STPA can be applied to an autonomous driving function on a car (adaptive cruise control (ACC)). A control structure diagram of the ACC system is used for the analysis (see Figure 2). It should be noted that as well as sensors, actuators and control modules, the diagram also includes the driver interactions.

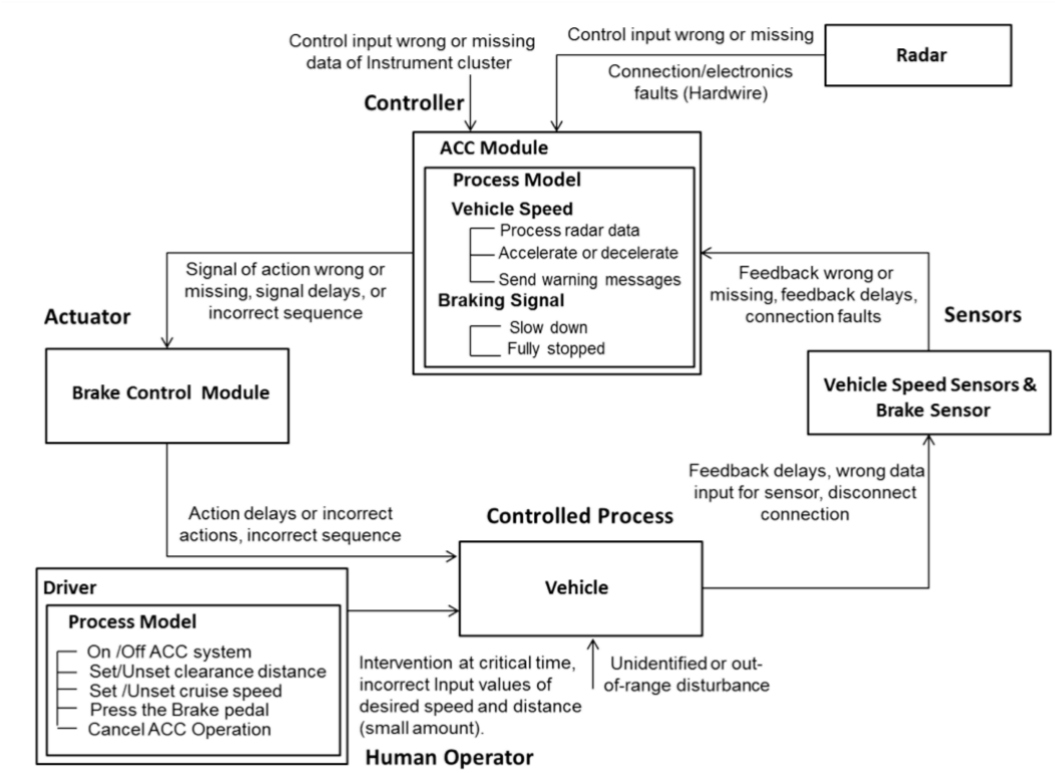

Figure 2 – A control structure diagram for ACC (taken from [6])

Deviations in the safety control actions are then considered. For example for the safety control action of providing radar data, four types of hazardous control actions are identified:

- Not Given: Radar sensor does not provide the relative speed and distance of objects ahead of vehicle.
- Given Incorrectly: Radar sensor provides incorrect data of target vehicle speed.
- Wrong Timing or Order: The data of radar sensor comes too late when the distance to a forward vehicle is too close.
- Stopped too soon or applied too long: Radar sensor is stopped too soon that the ACC module does not get the relative data signal of target vehicle.

In order to identify causal scenarios for the hazardous control actions the control structure is augmented with process models for each component. This last step of the STPA process is the most challenging, relying heavily on engineering domain knowledge and experience. There is currently no systematic method defined for doing this analysis. There is also currently no systematic way of considering the relationship between the controller state and the control actions. This could be achieved through integrating analysis of state machines, but this is not currently common practice.

A further example of the application of STPA is provided in [4] for the ALC system. This analysis identifies over 1,000 unique causal factors that could lead to unsafe control actions, illustrating the complexity of the approach.

Summary of approach

1. Produce a functional description of vehicle behaviour to be analysed and identify applicable operating scenarios
2. Identify possible deviations in vehicle behaviour through application of a systematic analysis method
3. Analyse the deviations in the relevant operating scenarios to determine which could result in a vehicle hazard

References

- [1] ISO, 2011, 'Road vehicles - Functional safety' (ISO 26262) , ISO, Geneva, Switzerland.
- [2] International Electrotechnical Commission., 2001. Hazard and operability studies (HAZOP Studies) - Application guide, Edition 1.0. (IEC 61882-2001).
- [3] Bagschik, G., Reschka, A., Stolte, T. and Maurer, M., 2016, June. Identification of potential hazardous events for an Unmanned Protective Vehicle. In 2016 IEEE Intelligent Vehicles Symposium (IV) (pp. 691-697). IEEE.
- [4] Becker, C., Yount, L., Rosen-Levy, S., & Brewer, J. (2018, August). Functional safety assessment of an automated lane centering system (Report No. DOT HS 812 573). Washington, DC: National Highway Traffic Safety Administration.
- [5] Leveson, N., 2011. Engineering a safer world: Systems thinking applied to safety. MIT press.
- [6] Abdulkhaleq, A. and Wagner, S., 2013. Experiences with applying STPA to software-intensive systems in the automotive domain. 2013 STAMP Conference, Boston, USA.