Risk Monitoring and Mitigation for Automated Vehicles: A Model Predictive Control Perspective

Kailin Tong, Fengwei Guo, Selim Solmaz, Martin Steinberger and Martin Horn

www.v2c2.at

Assoc. Prof. Dr. Selim Solmaz
Group Leader Control Systems

IEEE IAVVC, 16-18 Oct, 2023

17/10/2023

Motivation

Why is Risk-aware Planning important?

- Highly automated driving (HAD) requires the capability to detect and handle hazardous events to ensure safety and bring the vehicle to a safe state (SAE J3016, UN-ECE ALKS).

- Planning is faced with real-world problems, e.g., perception system degradation/failure, or the change of intention of another vehicle.

Why do we need a Monitoring Device (MonDev)

- Real-time or runtime monitoring device/functionality to supervise the automated driving system status to initiate a Minimum Risk Maneuver.

Why do we need MRM

- Plan a trajectory to bring the vehicle to a safe state to minimize the overall risk at an acceptable level

The Tesla accident due to perception failure

Tesla cannot correctly label a carriage

https://www.linkedin.com/posts/activity-6965342247278018560-d1yV/?utm_source=share&utm_medium=member_desktop

Related work

- Monitoring Device implementation and demonstration [1]
- Minimal Risk Maneuver

- However, a theoretical framework for dealing with autonomous vehicle hazards has rarely been presented. This study suggests a risk modeling method inspired by ideas from control theory and introduces a Model Predictive Control (MPC) Framework to deal with risks in general. [2]

[1] YouTube. (2023a). *VIRTUAL VEHICLE ArchitectECA2030 Demonstrator*. *YouTube*. Retrieved October 2, 2023, from https://www.youtube.com/watch?v=67ldtb56W-4.
[2] Tong, K., Solmaz, S., & Horn, M. (2022, October). A Search-based Motion Planner Utilizing a Monitoring Functionality for Initiating Minimal Risk Maneuvers. In 2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC) (pp. 4048-4055). IEEE.

Problem Formulation

A. *Problem Formulation*

For the system dynamics, we use the common state transition model.

$$\mathbf{x}_{k+1} = F(\mathbf{x}_k, \mathbf{u}_k) \tag{1}$$

where $\mathbf{x}_k \in \mathbb{R}^n$ is a state vector, $\mathbf{u_k} \in \mathbb{R}^m$ is the control input.

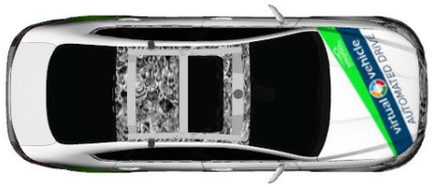
To evaluate the risk of a hazard, we use the most common definition of risk: risk is the probability times severity [1]. We first define the severity s_k of a hazard at time step k. Its severity model is given by

$$s_{k+1} = S(\mathbf{x}_{k+1}, s_k) \tag{2}$$

We denote a hazard as H for the Automated Driving System. The state is either h, meaning occurrence of the hazard, or $\bar{h}$, meaning no occurrence of the hazard. The observation monitoring the hazard at the k-th step is denoted as $\mathbf{z}_k$. We define the probability of hazard conditioned on observations as $p_k = p(h|\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_k)$. Finally, we have the risk model:

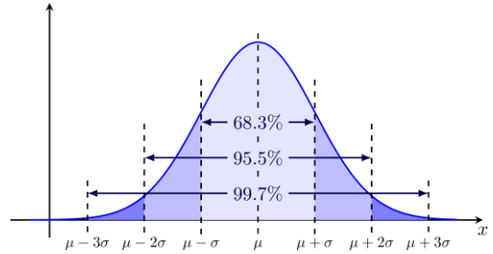
$$r_k = s_k \cdot p_k \tag{3}$$

State transition model



Severity model

Probability model



Risk model

Definitions

Definition 1 (Safe Set) *The state vector $\mathbf{x}_k$ in a bounded space $\mathbb{X} \subset \mathbb{R}^n$ and the hazard-related observation $\mathbf{z_k}$ in a bounded space $\mathbb{Z} \subset \mathbb{R}^m$ span a space $\chi \subset \mathbb{R}^{n+m}$. α is an acceptable risk level, which defines the boundary of the Safe Set $\chi^\alpha \subset \chi$. The Safe Set represents the normal operation of the Automated Driving System (ADS).*

The risk for a hazard is monitored during the operation of the ADS. As the standard ISO 26262 requires, a risk mitigation system keeps the AV safe if a hazard occurs.

Definition 2 (Stability of Risk Mitigation): *The risk of a hazard can be brought back to the Safe Set in $[t_0, t_{f1}]$ if the risk is greater than an acceptable risk level α, as shown in Fig. 1 .*

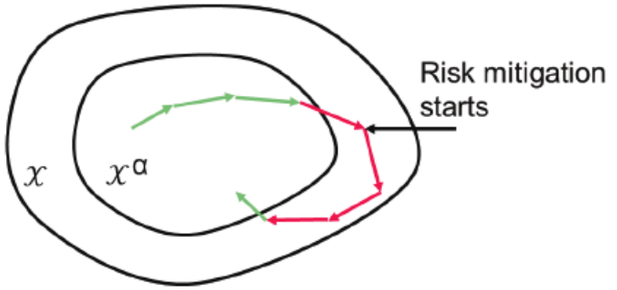
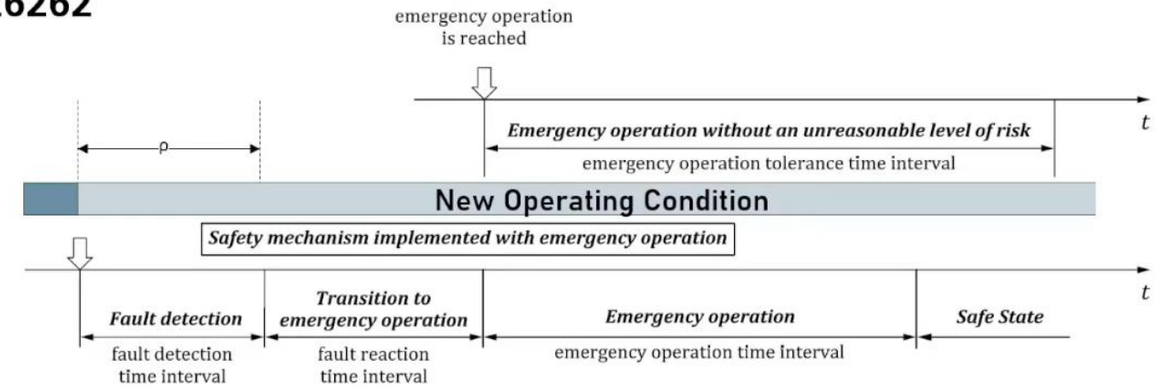


Fig. 1: Illustration of stability of risk mitigation.

Transiently Unsafe

ISO 26262



[3] A. Salvi, G. Weiss, M. Trapp, F. Oboril and C. Buerkle, "Safety Implications of Runtime Adaptation to Changing Operating Conditions," *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*, Macau, China, 2022, pp. 2444-2449, doi: 10.1109/ITSC55140.2022.9922192.

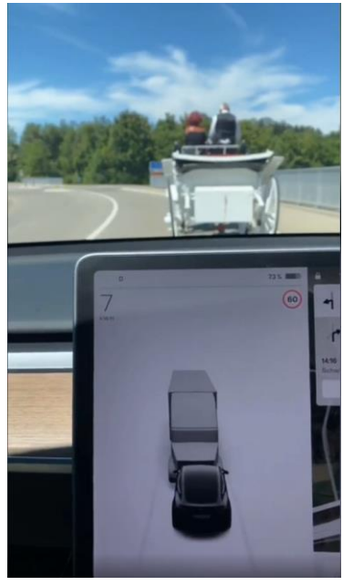
Definitions

Definition 3 (Controllability of a Hazard): *If a hazard occurs, there exists a vector of* $\mathbf{u}_k$ *that can bring the risk of a hazard to the Safe Set in* $[t_0, t_{f2}]$*, and the hazard is controllable.* t_{f2} *is the deadline for tolerating the existence of a hazard.*

Definition 4 (Observability of a Hazard): *If there exists a vector of Observations* $\mathbf{z}_k$ *measured by the ADS which can estimate the probability of a hazard in* $[t_0, t_{f3}]$*, the hazard is observable.* t_{f3} *is the deadline for detecting the existence of a hazard.*

Controllable

Not Controllable



Observable?

Observable?

Software & Hardware Architecture

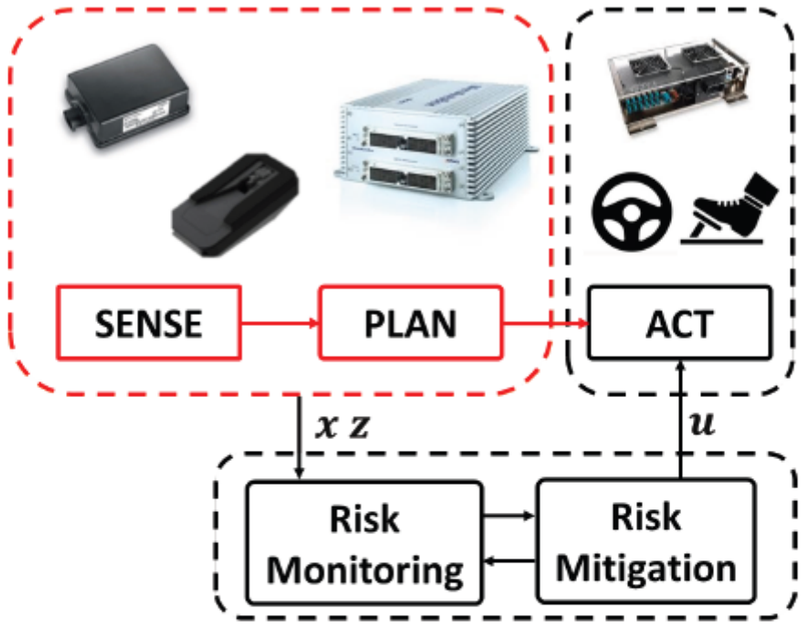


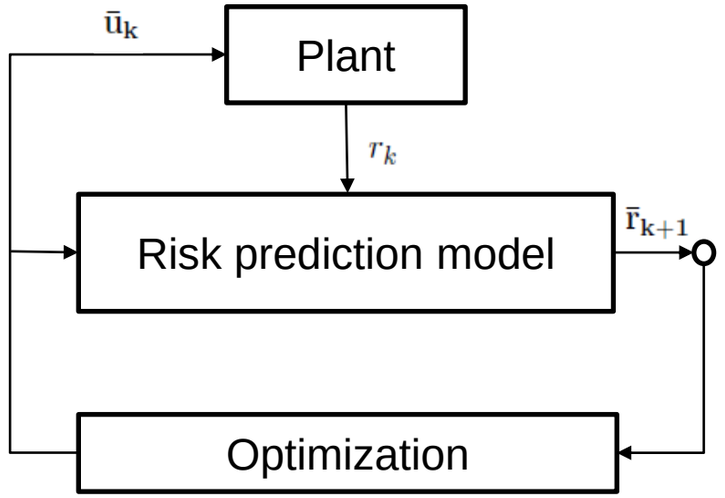
Fig. 2: Illustration of the proposed architecture for our AD demonstrator. The dashed block denotes hardware, while the solid block denotes software. We monitor the hardware and software of SENSE and PLAN, which are marked with red color. The Risk Monitoring and Mitigation (RMM) Module on the lower side is an additional software block to monitor and mitigate the risk.

MPC Framework

The proposed RMM module monitors the risk online and triggers risk mitigation. For any nominal planner, the Risk Mitigation Module is a supervisory planner that enforces the risk to return to χ^{α} if a hazard happens (its risk is higher than the acceptable level α). To simplify our formulation, we let $\mathbf{u}_k \in R$ and $\bar{\mathbf{u}}_k$ be the control until the end of the prediction horizon N_p. Motivated by Safety Barrier Functions [7], we have the following quadratic programming (QP) problem with hard constraints:

$$\begin{aligned} \mathbf{u}^* = \underset{\bar{\mathbf{u}}_k \in \mathbb{U}}{\arg\min} \quad & ||\bar{\mathbf{u}}_k - \mathbf{u}_0||^2_{\mathbf{Q}} \\ s.t. \quad & r_i \leq \alpha, \quad k + N_d \leq i \leq k + N_p \end{aligned} \tag{4}$$

where $\mathbf{u}_0$ is a vector of length N_p where each element is the output from a nominal planner at time step k. N_d is the deadline for bringing the risk less than α. $\bar{\mathbf{u}}_k = [u_k, u_{k+1}, \ldots, u_{k+N_c}, \ldots, u_{k+N_p-1}]^T$, where N_c is the control horizon. If N_c is smaller than N_p, after $k + N_c$ step, the control is same as u_{k+N_c}.



N_p	prediction horizon
N_c	control horizon
N_d	deadline
Q	weighting matrix
$\bar{r}_k$	predicted risk vector at step k
r_k	risk at step k
u_0	nominal output vector
$\bar{u}_k$	control vector at step k

Use Case 1: Hardware Hazard

Camera Offline:

- Considering an automated vehicle is driving in a suburban area with ACC (Adaptive Cruise Control), we envisage acritical hazard where the camera is offline, and the vehicle has no redundant sensor configuration. Hence the ACC function cannot continue.
- **State transition model**

$$v_{k+1} = v_k + \Delta T u_k$$

- **Severity model**

$$s_k = (v_k - v_s)^2.$$

- **Probability model**

$$p_k = \begin{cases} 1, & \text{if } z_k = 1 \\ 0, & \text{otherwise} \end{cases}$$

MPC Problem

$$\mathbf{u}^* = \arg\min_{\bar{\mathbf{u}}_\mathbf{k}} \quad ||\bar{\mathbf{u}}_\mathbf{k} - \mathbf{u_0}||^2_\mathbf{Q}$$

$$s.t. \quad -\sqrt{\alpha} + v_s - v_k \leq \sum_{j=0}^{i-1} \Delta T u_{k+j} \leq \sqrt{\alpha} + v_s - v_k,$$

$$i = N_d, N_{d+1}, \ldots, N_p,$$

$$0 \leq v_{k+i} \leq v_{max}, i = 1, 2, \ldots, N_p,$$

$$u_i \in [u_{min}, u_{max}], i = k, k+1, \ldots, k+N_p-1, \quad (9)$$

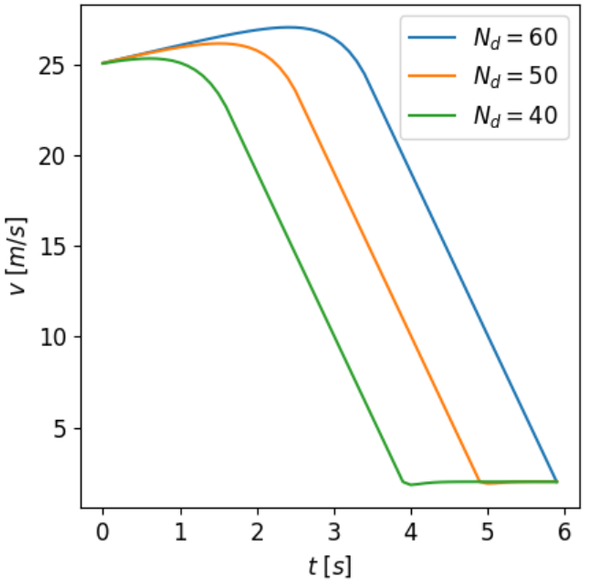
N_p	prediction horizon
N_d	deadline
α	risk threshold
Q	weighting matrix
v_k	velocity at step k
p_k	probability at step k
s_k	severity at step k
u_0	nominal output vector
$\bar{u}_k$	control (acceleration) vector at step k

Use Case 1: Hardware Hazard

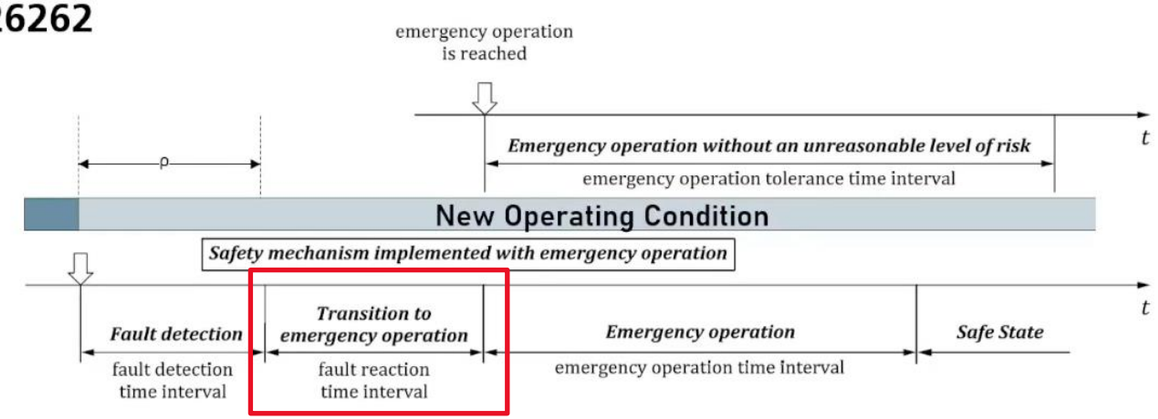
The effect of Deadline N_d

- Smaller Nd → quicker response (shorter transition time)
- Larger Nd → slower response (longer transition time)

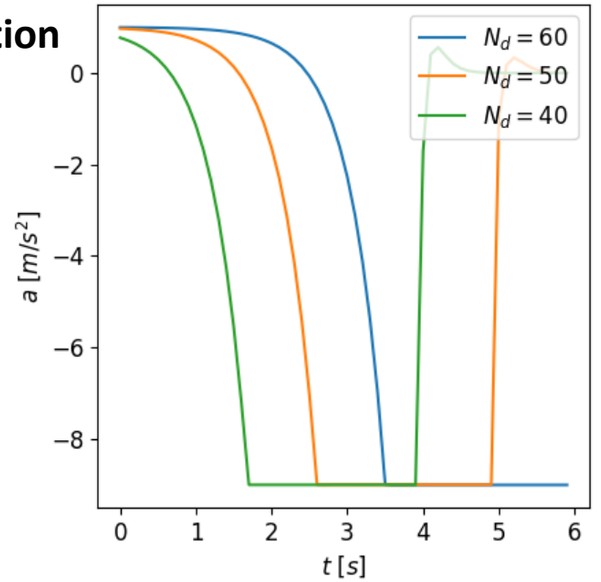
Velocity



ISO 26262



Acceleration



Use Case 2: Software Hazard

Error of Perception Algorithm

- ADS faces significant challenges in accurately perceiving the surrounding environment, including identifying and continuously tracking surrounding objects. A failure of the perception module often results in severe traffic accidents

Confirmation problem in object tracking

- Verify tracking and triggering MRM if object tracking fails

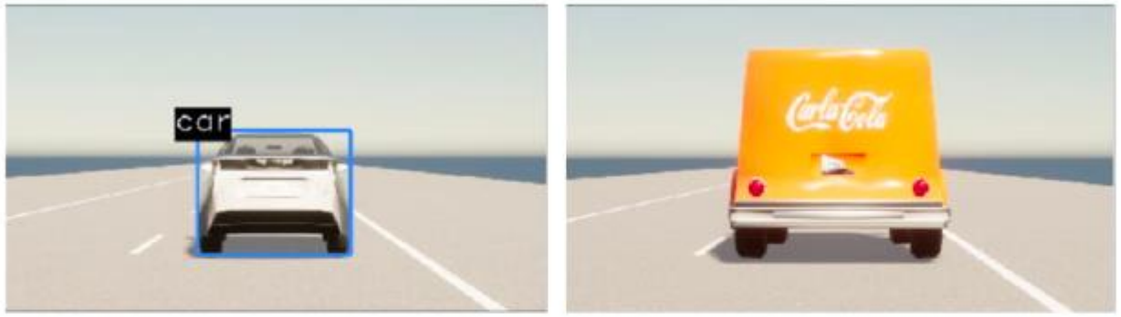
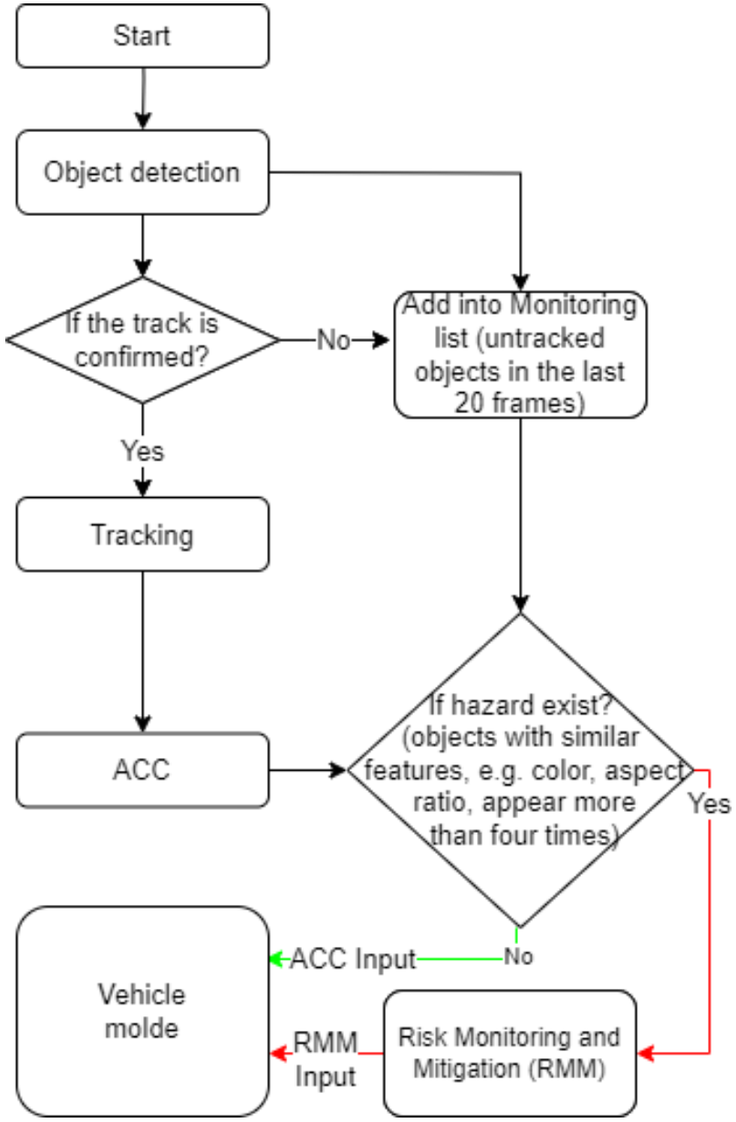


Fig. 4: Ground Vehicle Targets: Toyota Prius (left) and CarlaCola (right). The perception system can track the Toyota Prius precisely. However, it occasionally loses track of the CarlaCola.



Use Case 2: Software Hazard

- **State transition model**

$$x_{k+1} = Ax_k + Bu_k,$$

$$\text{where } x_k = \begin{bmatrix} d_k \\ v_k \end{bmatrix}, A = \begin{bmatrix} 1 & \Delta T \\ 0 & 1 \end{bmatrix}, B = \begin{bmatrix} \frac{1}{2}\Delta T^2 \\ \Delta T \end{bmatrix} \text{ and}$$

$$y_k = Cx_k,$$

- **Severity model**

$$s_k = v_k$$

- **Probability model**

$$p_k = \begin{cases} 1, & \text{if a perception error is detected} \\ 0, & \text{otherwise} \end{cases}$$

MPC Problem

$$\begin{aligned} \mathbf{u}^* = \arg\min_{\bar{\mathbf{u}}_\mathbf{k}} \quad & ||\bar{\mathbf{u}}_\mathbf{k} - \mathbf{u_0}||^2_\mathbf{Q} \\ s.t. \quad & v_{k+i} \leq \alpha, i = N_d, N_{d+1}, \ldots, N_p \\ & 0 \leq v_{k+i} \leq v_{max}, i = 1, \ldots, N_p \\ & u_i \in [u_{min}, u_{max}], i = k, k+1, \ldots, k+N_p-1 \end{aligned} \tag{15}$$

N_p	prediction horizon
N_d	deadline
α	risk threshold
Q	weighting matrix
v_k	velocity at step k
p_k	probability at step k
s_k	severity at step k
u_0	nominal output vector
$\bar{u}_k$	control (acceleration) vector at step k

Use Case 2: Software Hazard

Simulation Environment

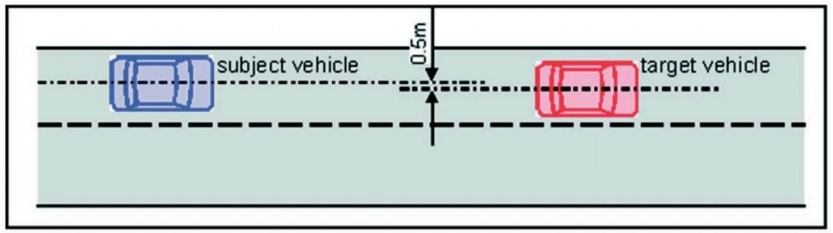
Carla Simulator + Vision-based ACC based on Autoware.AI

Scenarios

TABLE I: Euro NCAP ACC Car-to-Car test scenarios with Stationary and Moving Target (straight roads) [12]

Scenarios	Vehicle under Test	Global Vehicle Target
Car-to-Car Rear Stationary (CCRs)	70, 80, 90, 100, 110, 120, 130 km/h	0 km/h
Car-to-Car Rear Moving (CCRm)	80, 90, 100, 110, 120, 130 km/h	20 km/h
	80, 90, 100, 110, 120, 130 km/h	60 km/h

[4} Euro NCAP, “Euro NCAP Assisted Driving - Highway Assist Systems Test and Assessment Protocol v1.0,” Euro NCAP, Tech. Rep., September 2020. [Online]. Available: https://cdn.euroncap.com/media/58813/euroncap-ad-test-and-assessment-protocol-v10.pdf



Automatic stop capability test ISO 15622:2018(E) (2018)

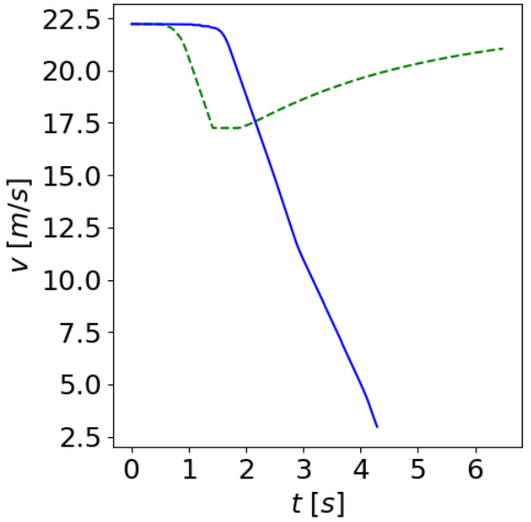
Use Case 2: Software Hazard

TABLE II: Simulation Results: Collision Avoidance Rate with respect to Euro NCAP ACC Car-to-Car test scenarios with Stationary and Moving Target (straight)

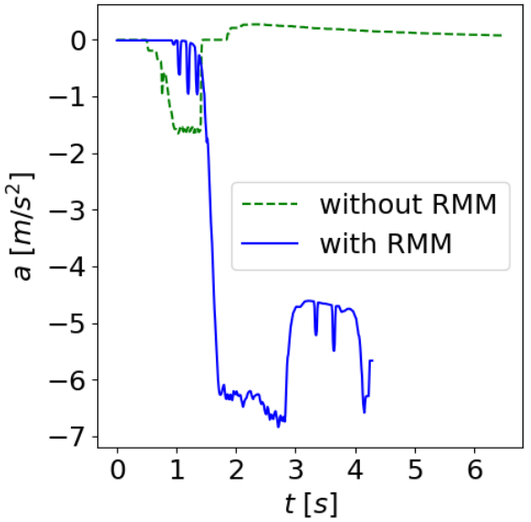
Scenario	Original	Hazardous, with RMM	Hazardous, without RMM
70 km/h vs. 0 km/h	100%	60%	0%
80 km/h vs. 0 km/h	100%	70%	0%
90 km/h vs. 0 km/h	100%	40%	0%
100 km/h vs. 0 km/h	100%	40%	0%
110 km/h vs. 0 km/h	0%	0%	0%
120 km/h vs. 0 km/h	0%	0%	0%
130 km/h vs. 0 km/h	0%	0%	0%
80 km/h vs. 20 km/h	100%	70%	0%
90 km/h vs. 20 km/h	100%	60%	0%
100 km/h vs. 20 km/h	100%	50%	0%
110 km/h vs. 20 km/h	100%	50%	0%
120 km/h vs. 20 km/h	100%	40%	0%
130 km/h vs. 20 km/h	0%	0%	0%
80 km/h vs. 60 km/h	100%	80%	0%
90 km/h vs. 60 km/h	100%	70%	0%
100 km/h vs. 60 km/h	100%	80%	0%
110 km/h vs. 60 km/h	100%	70%	0%
120 km/h vs. 60 km/h	100%	60%	0%
130 km/h vs. 60 km/h	100%	50%	0%
Average	78.94%	46.84%	0%

"**Hazardous**" refers to scenarios where the ego vehicle is following a CarlaCola car, and a tracking error occurs.
"**Original** " refers to scenarios where the ego vehicle is following another Carla car, and no tracking error occurs.

Velocity



Acceleration



Use Case 2: Hardware Hazard

Without RMM (Risk Monitoring and Mitigation)

Use Case 2: Hardware Hazard

With RMM (Risk Monitoring and Mitigation)

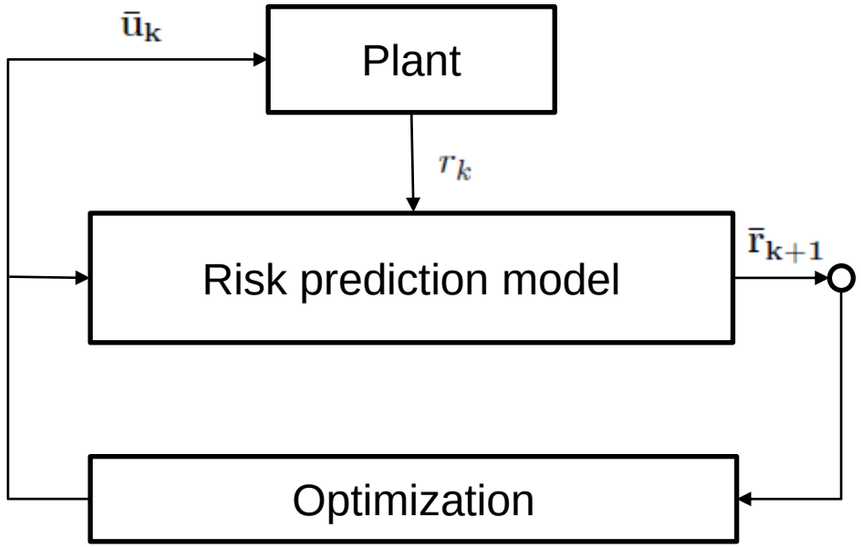
Key takeaways

Conclusion

- Bridging functional safety and control theory concepts by incorporating definitions such as risk mitigation stability, hazard controllability, and hazard observability.
- A novel Model Predictive Control (MPC) framework that addresses the handling of hazards.
- The effectiveness of the framework is demonstrated through two representative examples in simulation.

Outlook

- Extending the proposed framework to monitor and mitigate various hazards in more diverse scenarios.



ESRIUM
SAFE AND EFFICIENT ROADS

ArchitectECA2030

EUSPA
European Union Agency for the Space Programme

virtual vehicle
ENABLING **FUTURE VEHICLE** TECHNOLOGIES

Funding

The publication was written at Virtual Vehicle Research GmbH within the scope of the EU project ArchitectECA2030 and under grant agreement No 877539, and ESRIUM Project, which has received funding from the European Union Agency for the Space Programme under grant agreement No 101004181. The project is co-funded by grants from Germany, Netherlands, Czech Republic, Austria, Norway and - Electronic Component Systems for European Leadership Joint Undertaking (ECSEL JU). The authors would like to further acknowledge the financial support within COMET Competence Centers for Excellent Technologies from the Austrian Federal Ministry for Climate Action, the Austrian Federal Ministry for Labour and Economy, the Province of Styria (Dept. 12) and the Styrian Business Promotion Agency (SFG). The Austrian Research Promotion Agency (FFG) has been authorised for the programme management. The content of this paper reflects only the authors' view. Neither the European Commission nor the EUSPA is responsible for any use that may be made of the information it contains.

ESRIUM project has received funding from the European Union Agency for the Space Programme under the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004181.

THANK YOU!

Assoc. Prof. Dr. Selim Solmaz
Group Leader Control Systems

www.v2c2.at

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Arbeit und Wirtschaft
FFG Promoting Innovation.
Das Land Steiermark Wirtschaft, Tourismus, Wissenschaft und Forschung
SFG

Virtual Vehicle Research GmbH wird im Rahmen von COMET Competence Centers for Excellent Technologies durch BMK, BMAW, Land Steiermark und Steirische Wirtschaftsförderung (SFG) gefördert. Das Programm wird durch die FFG abgewickelt.