ESPLANADE

Towards Safety Analysis of Interactions Between Human Users and Automated Driving Systems

Fredrik Warg

Stig Ursing • Martin Kaalhus • Richard Wiik

RI. SE

semcon

Safety of Automated Driving Systems

- Need to argue that an ADS feature is sufficiently safe prior to release
- The automated driving system (ADS) must drive safely while in control
- **Safe interaction with human users (HU)**

Note: Terminology used mainly from SAE J3016 *"Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles"*

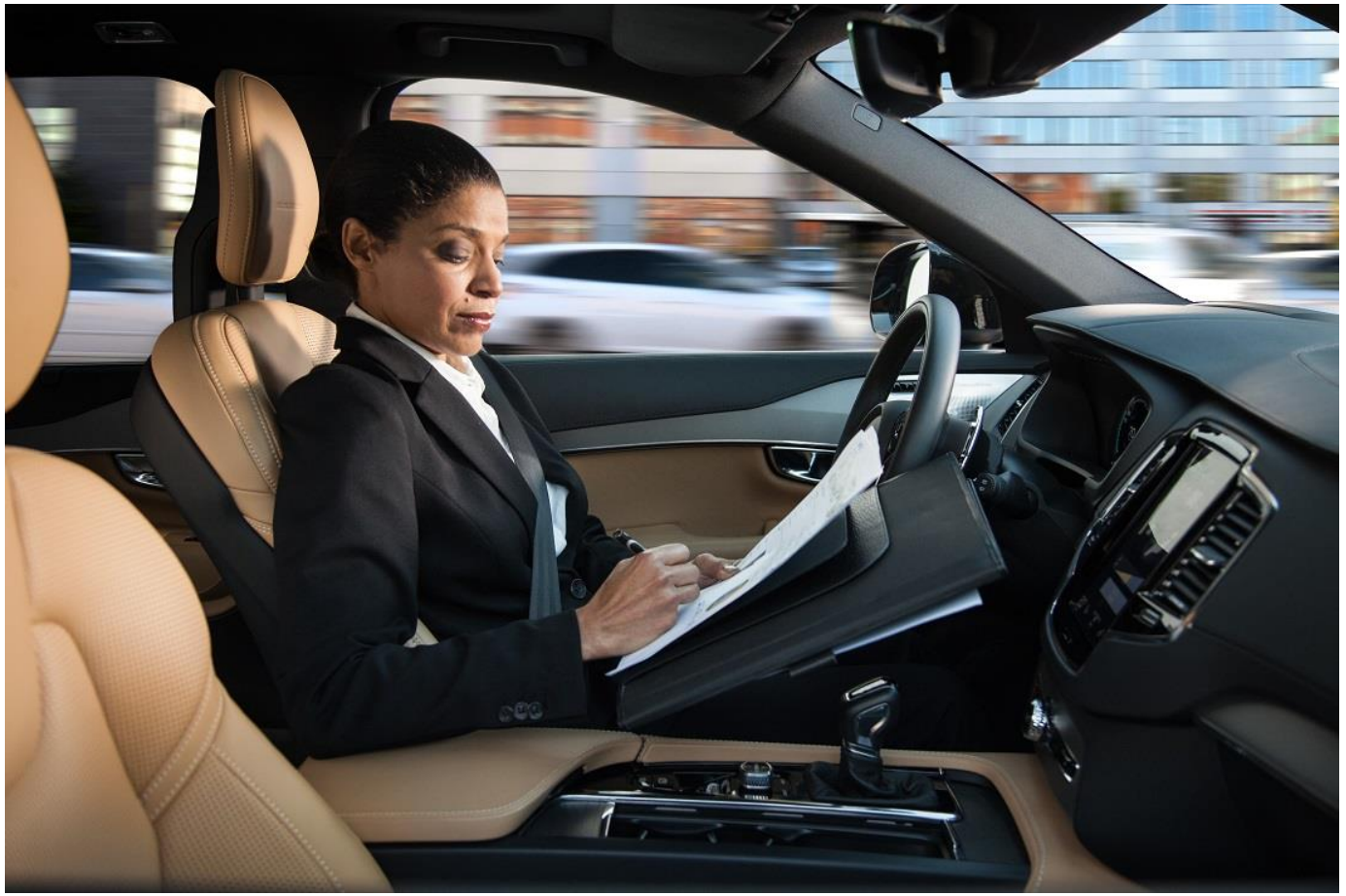
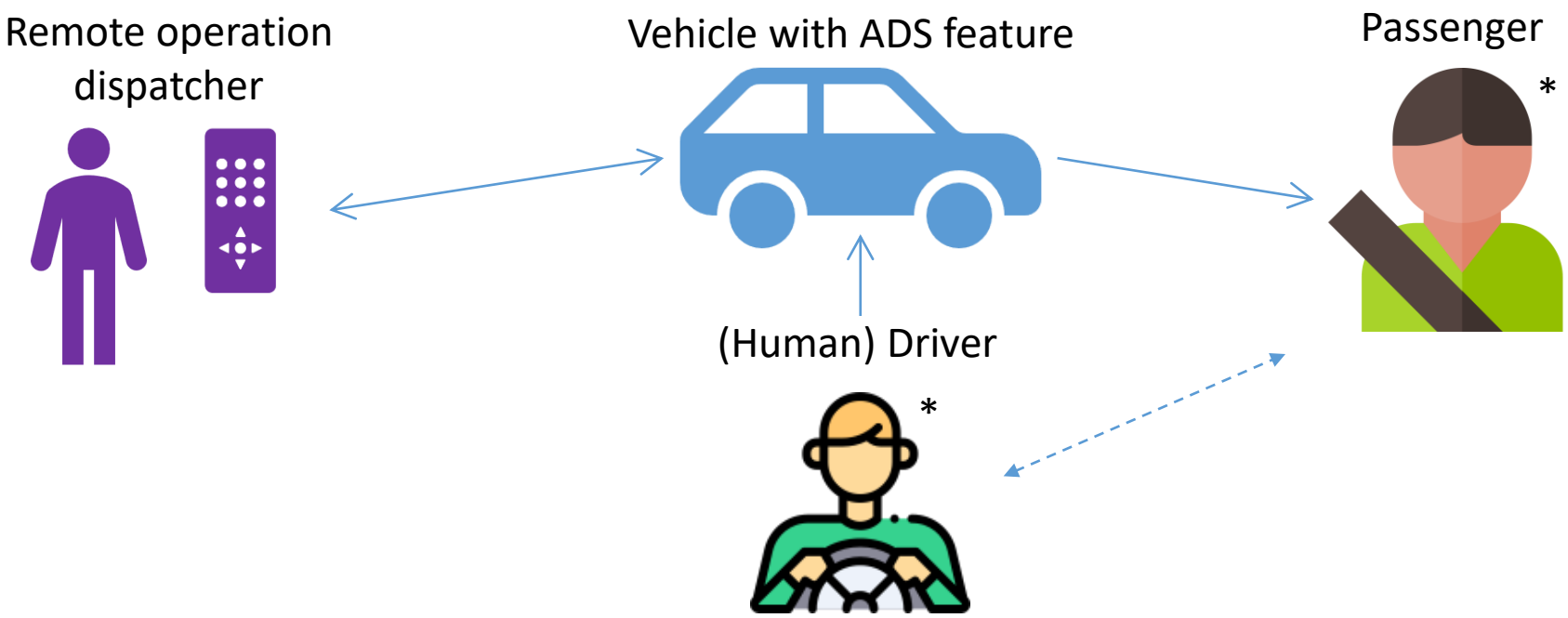


Photo: Volvo Cars

Transitions of control between ADS and HU



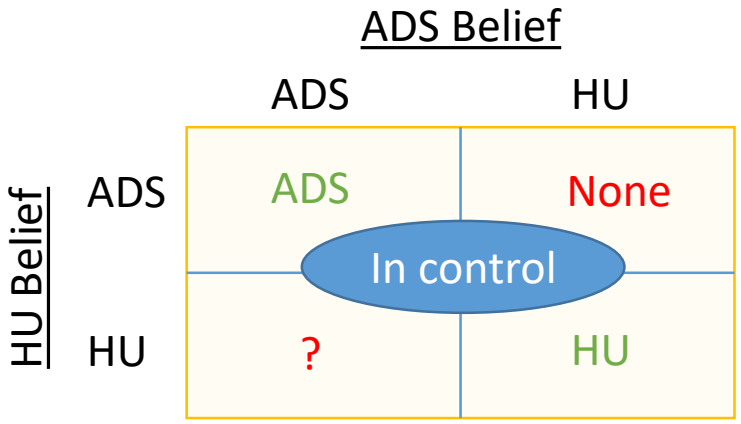
- Focus in this presentation: ***Transitions*** *of control between* ***human user*** *and* ***high driving automation*** *feature (SAE Level 4) in a moving vehicle.*

E.g. Highway pilot

*Icon made by Freepik from www.flaticon.com

Transition Hazards

Mode confusion



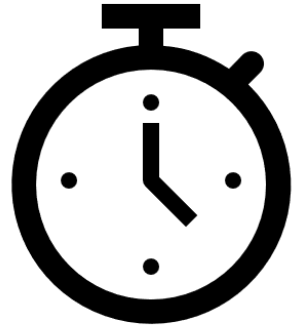
ADS and HU do not share belief of who is driving.

Unfair transition

Car photo created by yanalya - www.freepik.com

ADS or HU forced to take control when not prepared and able to drive.

Stuck in transition



ADS or HU unable to complete transition in time, impairing driving capability.

Ref: Johansson et al. ”Safe Transitions of Responsibility in Highly Automated Driving”, 2016 and ”*Safe Transitions Between a Driver and an Automated Driving System*”, 2017.

Safe Transitions

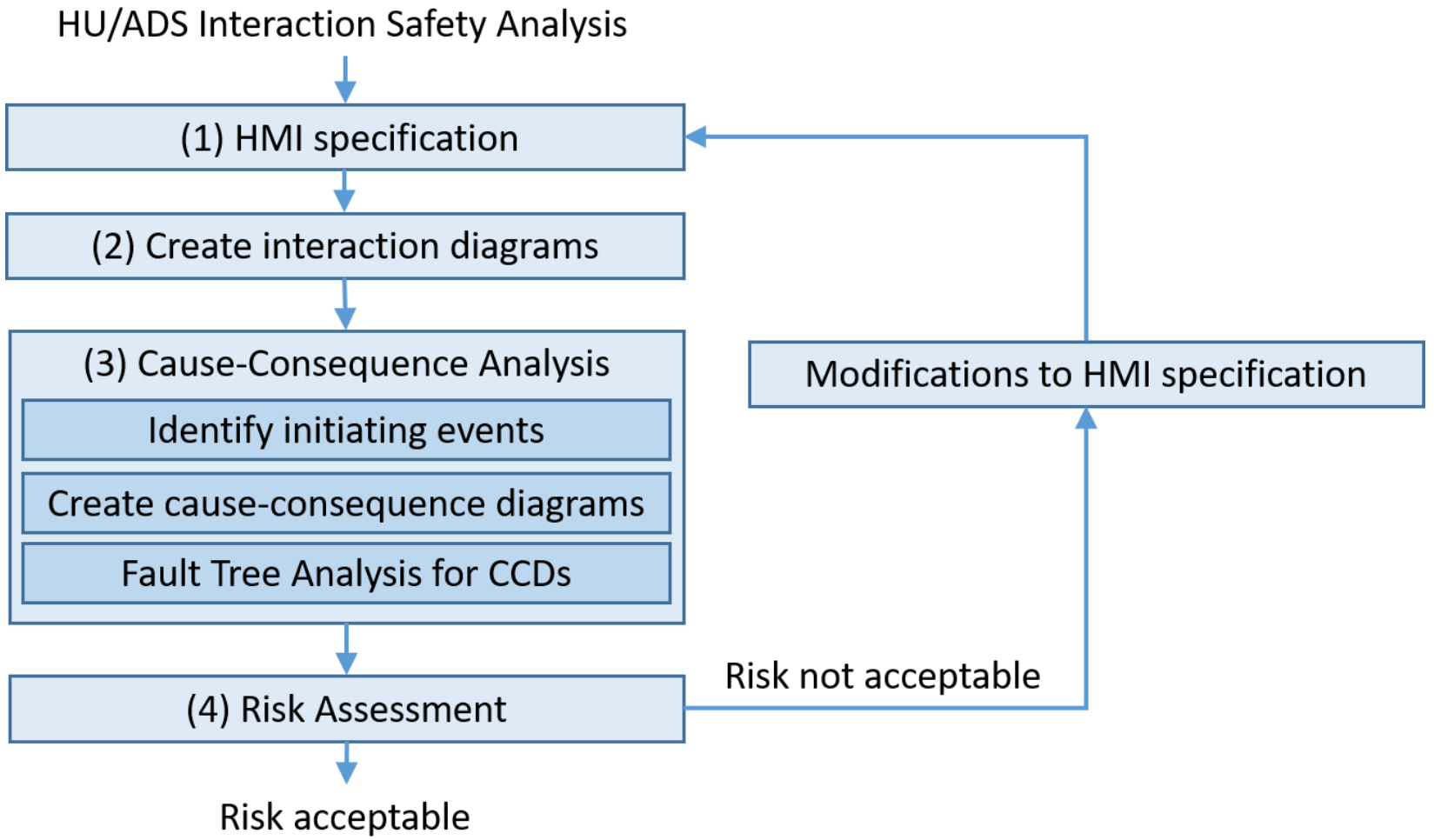
- Previous work:
  - Transition hazards
  - Principles for safe handover
  - Safety analysis for a transition protocol
- In this work:
  - Propose method to perform safety analysis combining practices from functional safety and human factors
  - Goal: Provide systematic analysis method for safety argumentation

TABLE I. SAFETY ANALYSIS OF TRANSITION PROTOCOL

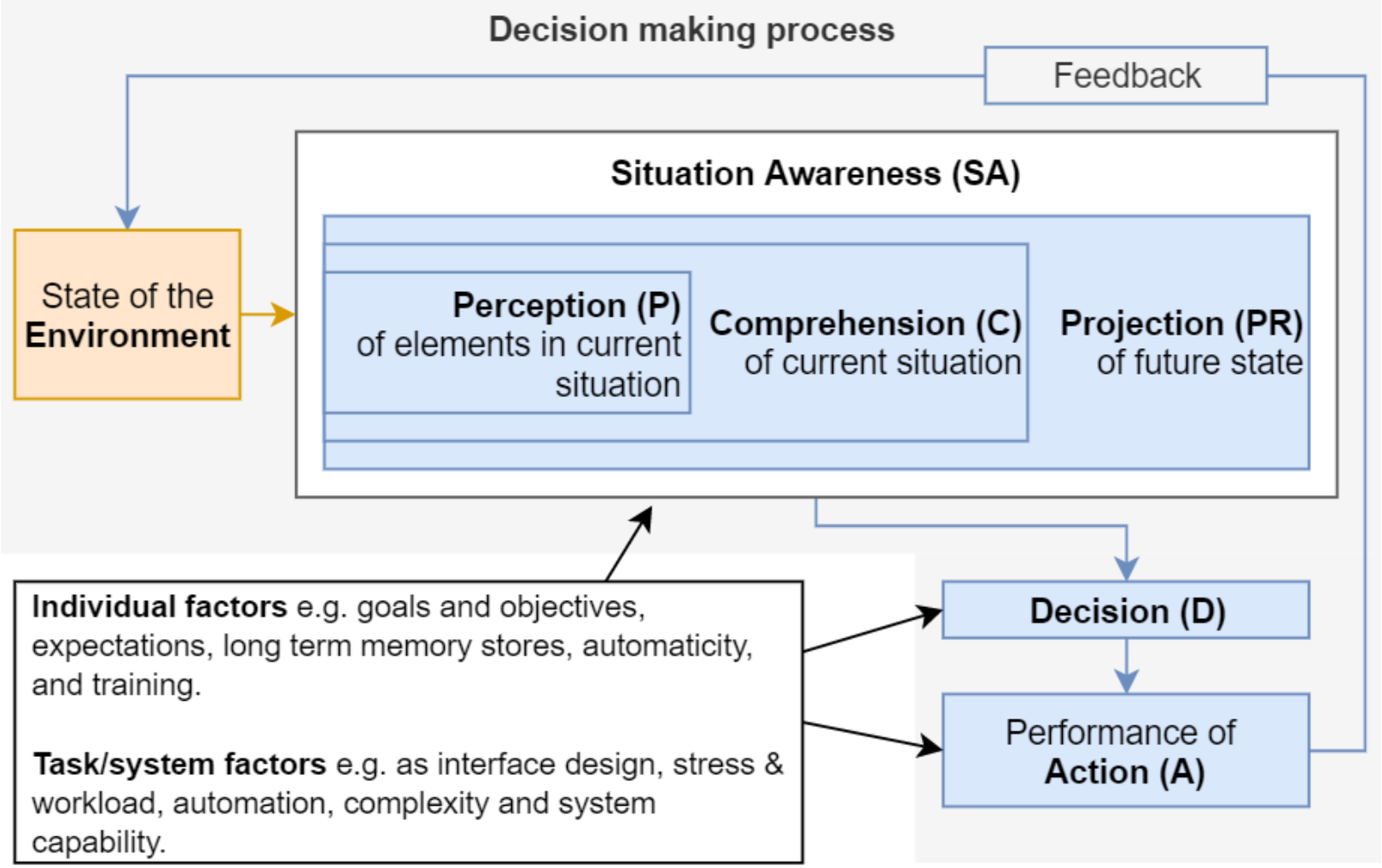
Protocol state	HMI failure	Driver mistake	Consequence	Safe/ Unsafe
MD - normal drive	Fault in lever lock	No	MD driver not trying to touch lever. Stay in MD.	Safe
MD - normal drive	Fault in lever lock	Driver changes lever position without asking for change first.	Unfair transition.	Unsafe
MD - normal drive	Fault in preference tell-tale	Any mistake or correct behaviour	MD cannot change locked lever. Stay in MD- normal drive.	Safe
MD - AD available	Fault in lever lock	No	MD driver not trying to touch lever. Stay in MD.	Safe
MD - AD available	Fault in lever lock	Driver changes lever position without asking for change first.	Unfair transition.	Unsafe
MD - AD available	Fault in preference tell-tale	No	Stay in MD	Safe
MD - AD available	Fault preference tell-tale	Driver ignores lack of availability	Transition sequence fulfilled. Change to AD.	Safe
MD - requested AD	Fault in push-button	Any mistake or correct behaviour	No Acknowledge by AD. Lever still locked. Stay in MD.	Safe

Source: Johansson et al. "*Safe Transitions Between a Driver and an Automated Driving System*", 2017.

Interaction Analysis Process

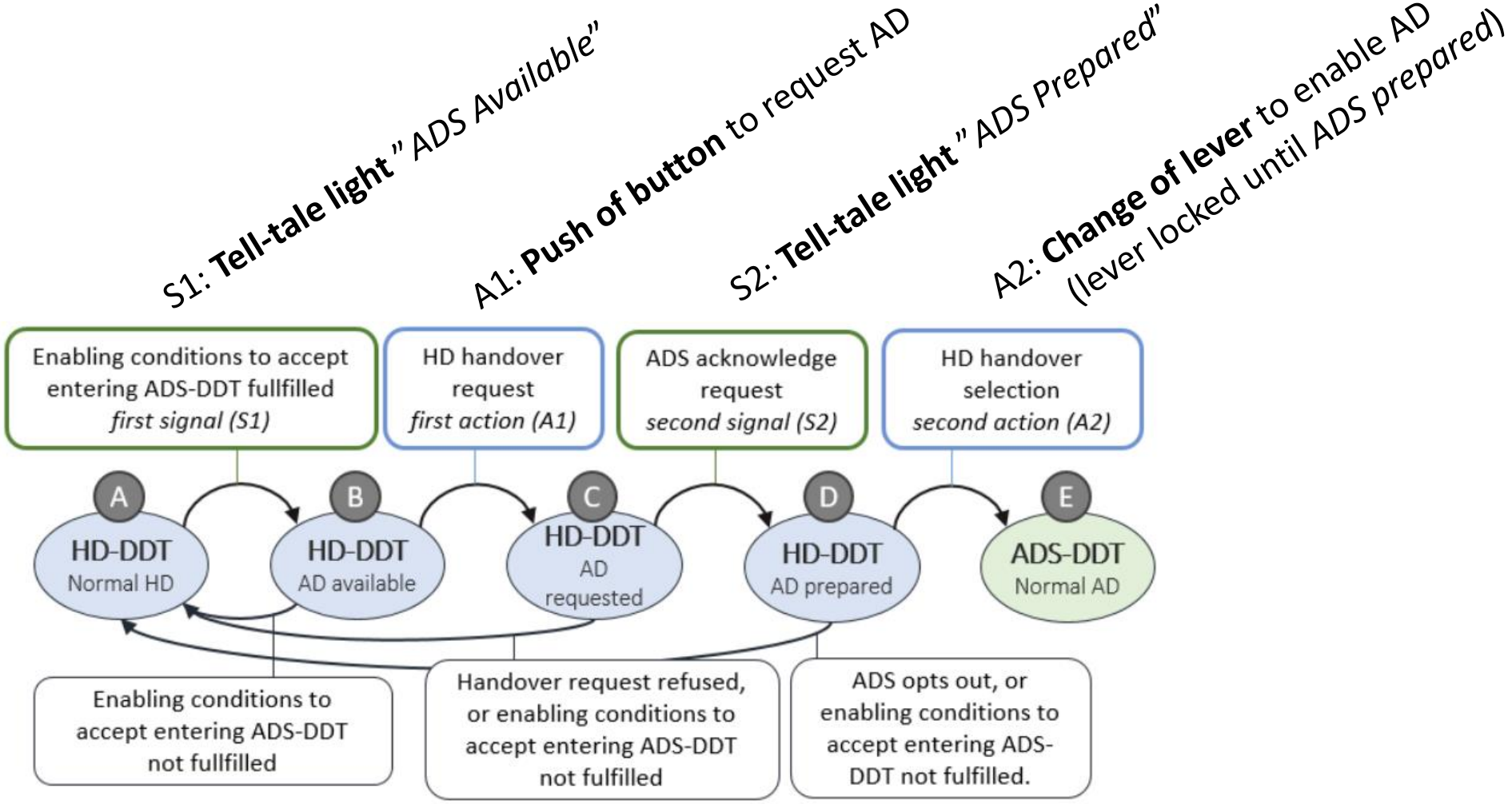


Human Performance Model



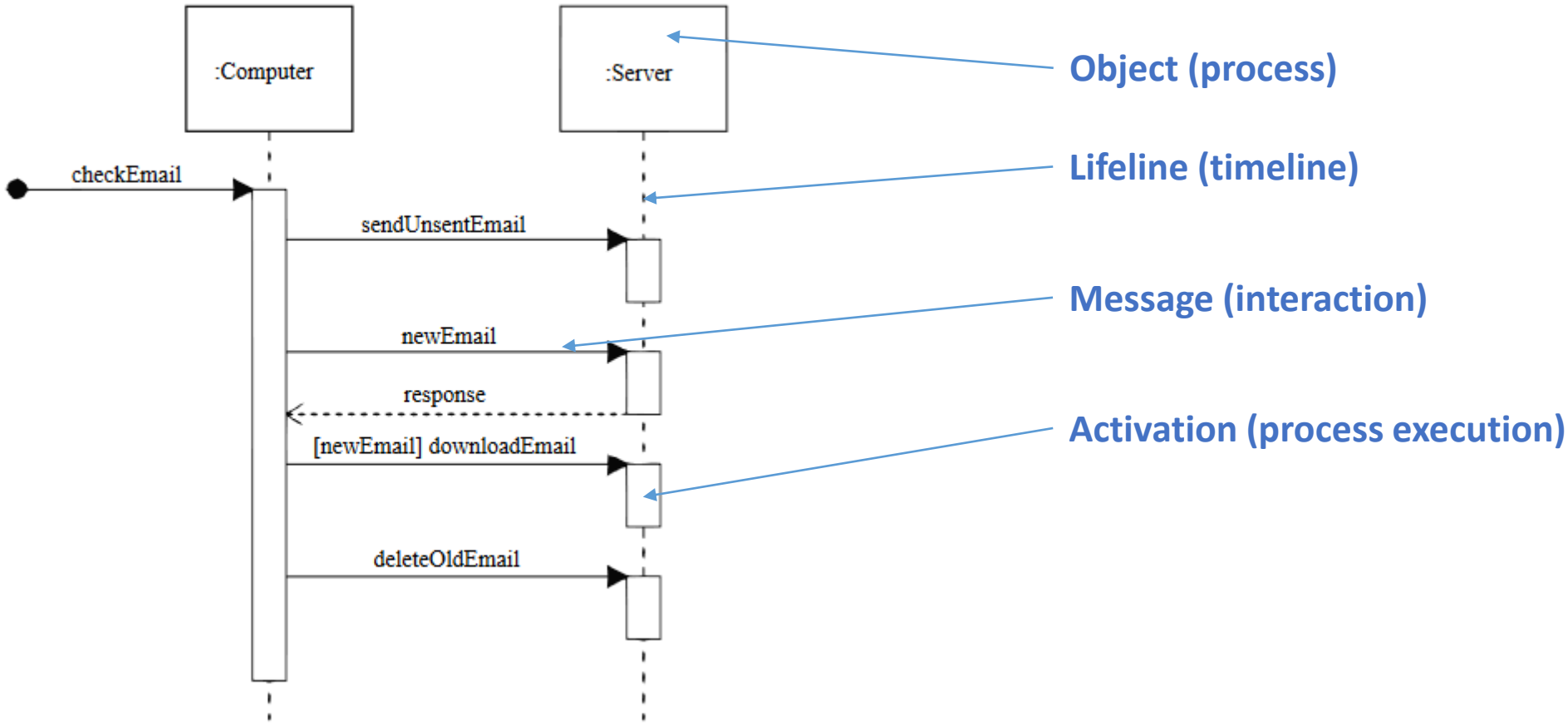
Ref: M. R. Endsley, "*Towards a theory of situation awareness in dynamic systems*", 1995.

HMI Specification – Illustrative Example



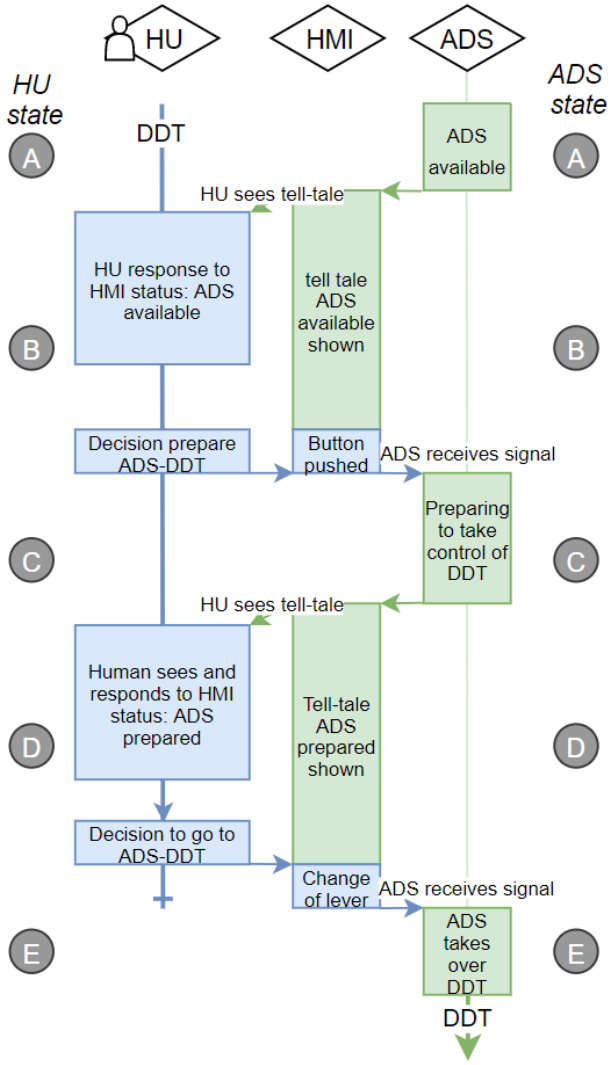
Example similar to: Johansson et al. ”*Safe Transitions Between a Driver and an Automated Driving System*”, 2017.

Sequence Diagrams (UML)

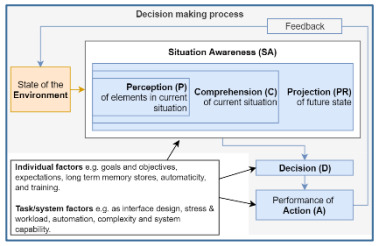


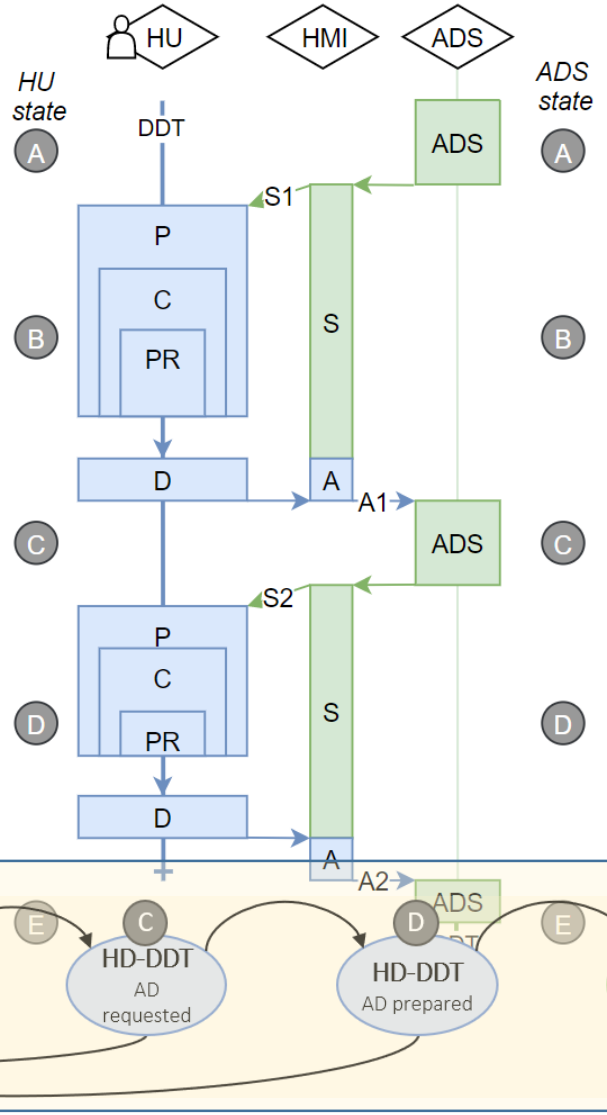
Source: Coupling_loss_graph.svg (https://commons.wikimedia.org/wiki/File:CheckEmail.svg), „CheckEmail“, https://creativecommons.org/licenses/by-sa/3.0/legalcode

Human-ADS Interaction Sequence Diagrams

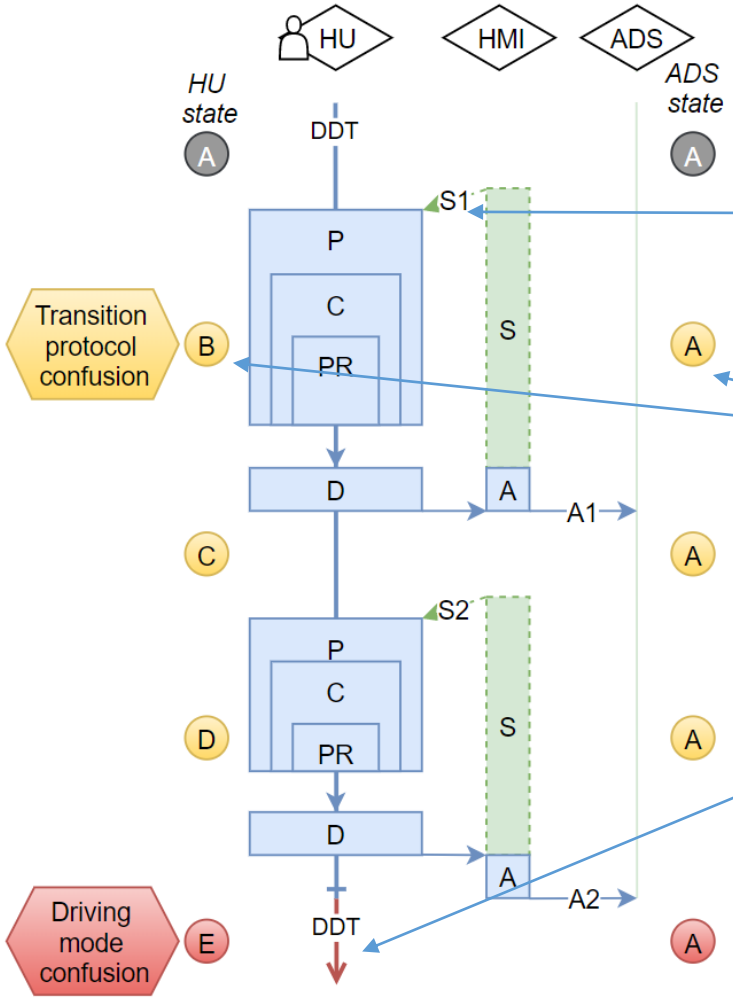


Human Performance Model





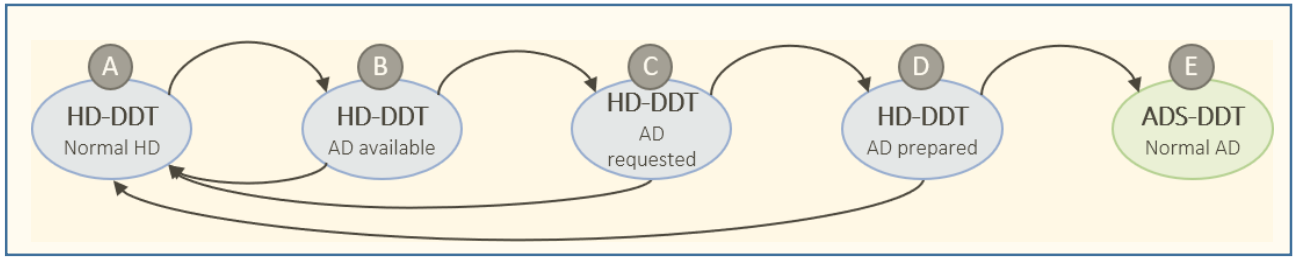
Interaction Sequence – Example with Hazard



Stimuli **S1** is perceived even though ADS has not initiated it.

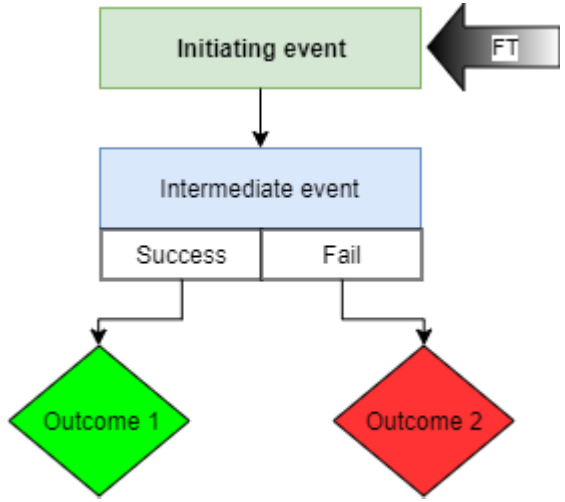
ADS and HU have different understanding of current protocol state → *Transition protocol confusion*.

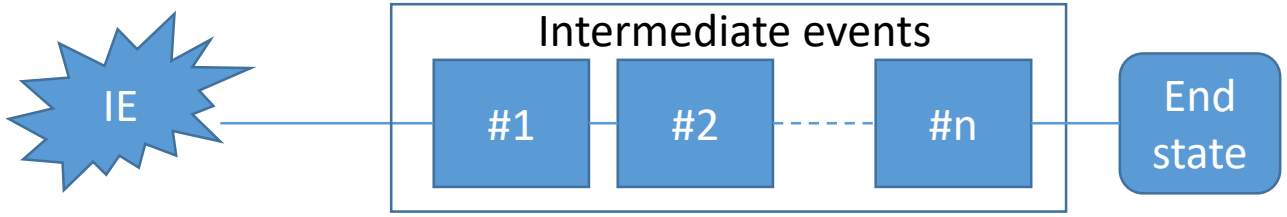
Transition protocol confusion may lead to a transition hazard, in this case **mode confusion**.



Cause-Consequence Analysis

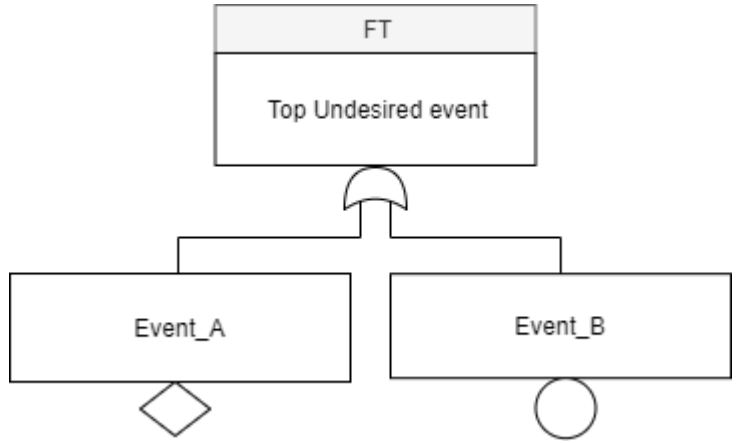
- Identify initiating events
- Identify intermediate events
- Build CCA diagram
- Use fault trees to determine how an event can fail





Undesired event that initiates start of accident sequence.

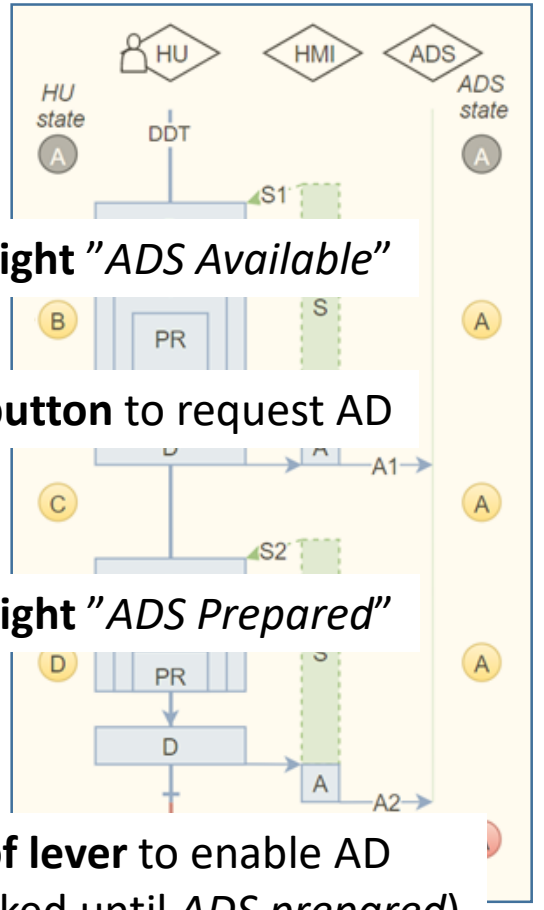
Failure/success events of safety measures designed to prevent IE from resulting in accident



Cause-Consequence Analysis: Our Example

- Identify initiating events

IE#	Initiating event	Explanation
IE1	S1 commission	S1 incorrectly provided
IE2	A1 commission	A1 performed without correct S1
IE3	S2 commission	S2 incorrectly provided
IE4	A2 commission	A2 performed without correct S2



S1: **Tell-tale light** ”*ADS Available*”

A1: **Push of button** to request AD

S2: **Tell-tale light** ”*ADS Prepared*”

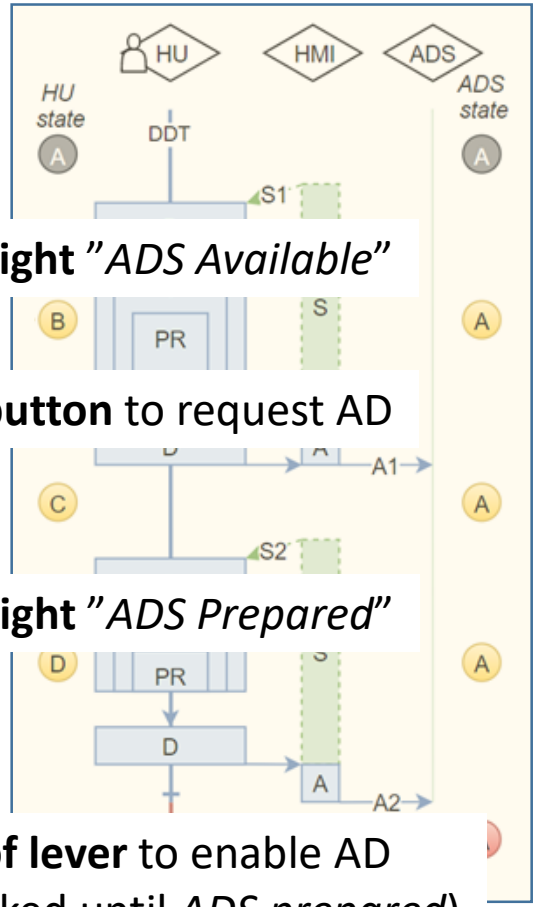
A2: **Change of lever** to enable AD
(lever locked until *ADS prepared*)

Cause-Consequence Analysis: Our Example

- Identify initiating events
- Identify intermediate events

IE#	Initiating event	Explanation
IE2	A1 commission	A1 performed without correct S1

Intermediate events
S2 performance
A2 performance



S1: **Tell-tale light** ”*ADS Available*”

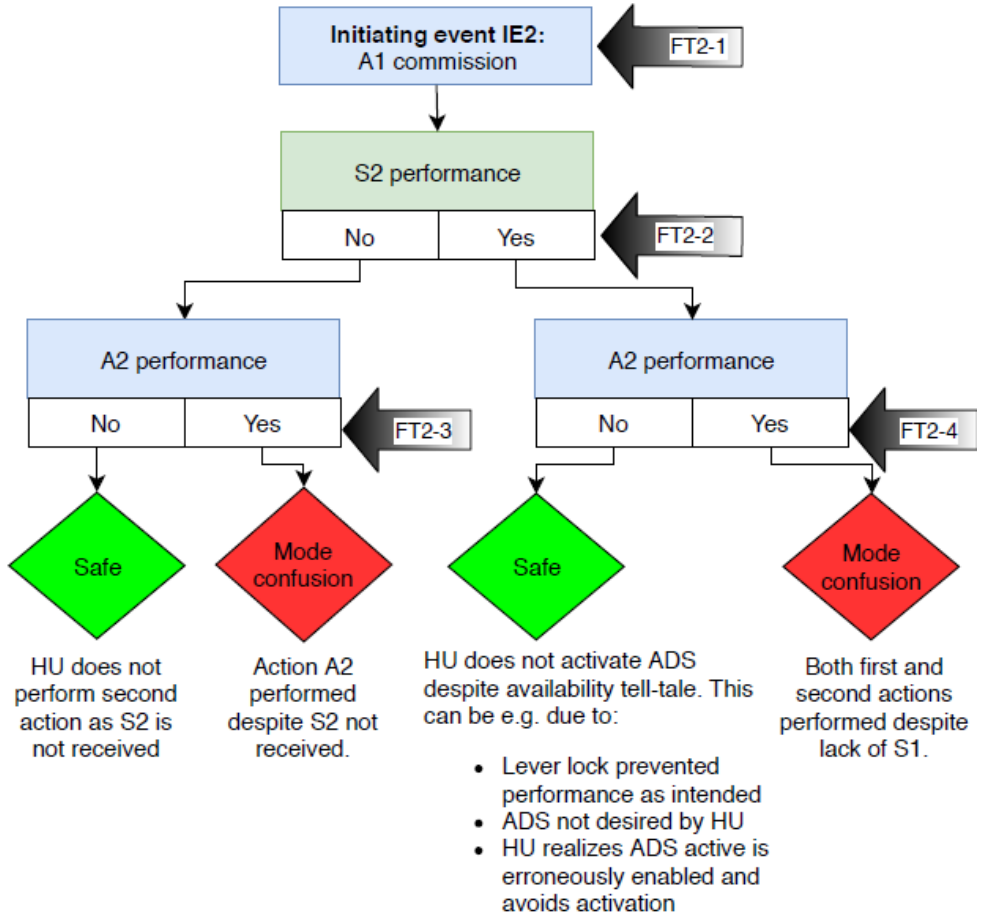
A1: **Push of button** to request AD

S2: **Tell-tale light** ”*ADS Prepared*”

A2: **Change of lever** to enable AD
(lever locked until *ADS prepared*)

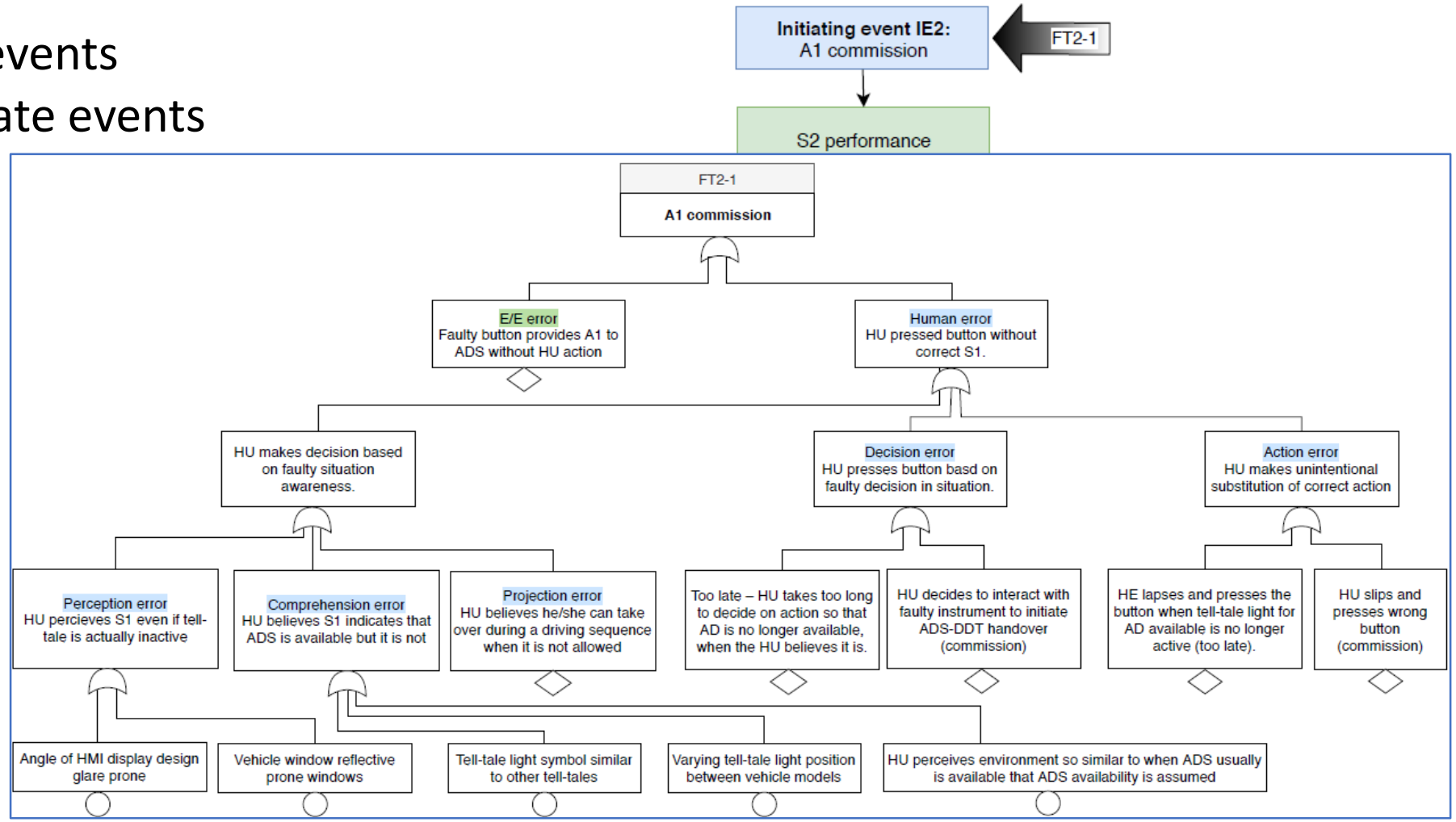
Cause-Consequence Analysis: Our Example

- Identify initiating events
- Identify intermediate events
- Build CCA diagram



Cause-Consequence Analysis: Our Example

- Identify initiating events
- Identify intermediate events
- Build CCA diagram
- Use fault trees to analyze how an event can fail



Risk Assessment and Risk Reduction

- CCA results used to improve HMI to reduce risk of transition hazards
  - Redesign
  - Adding safety measures
- How to do risk assessment? Further work needed.
- Iterative analysis/redesign until the HMI is sufficiently safe

In Summary

Conclusions

- Safety analysis of interactions between human users and ADS necessary for an ADS safety case
- We propose the use of an analysis method based on known techniques: sequence diagrams, cause-consequence analysis and the situation awareness model

Future Work

- Guidance for finding likely human errors in each of the categories (P/C/PR/D/A)
- How to capture risks of dependent or timing-related hazards?
- Interaction between driver capability and ODD and ADS feature specifications
- Alternatives to CCD, e.g. STPA
- Risk assessment method
  - Connection to ISO 26262

Also in our paper:

- Relation to standards in the automotive domain: ISO 26262 and ISO PAS 21448
- Discussion on terminology differences between functional safety and human factors domains

Thank you for listening!

Questions?

ESPLANADE

This research has been supported by Vinnova - Sweden's innovation agency, via the project ESPLANADE.

VINNOVA

APTIV · COMENTOR · KTH VETENSKAP OCH KONST · qamcom

RI. SE · semcon · SYSTEMITE · veoneer

VOLVO · VOLVO Volvo Group · ZENUITY Make it real.