

Hi-Drive – 1st Summer School, Porto Heli, Greece

Towards a Quantitative SOTIF Validation of Automated Driving Systems

Lina Putze

German Aerospace Center (DLR) e.V.

Institute of Systems Engineering for Future Mobility



How does the ISO 21448:2022 require or suggest performing a quantitative SOTIF validation?

How does the ISO 21448:2022 require or suggest performing a quantitative SOTIF validation?

To answer this, we...

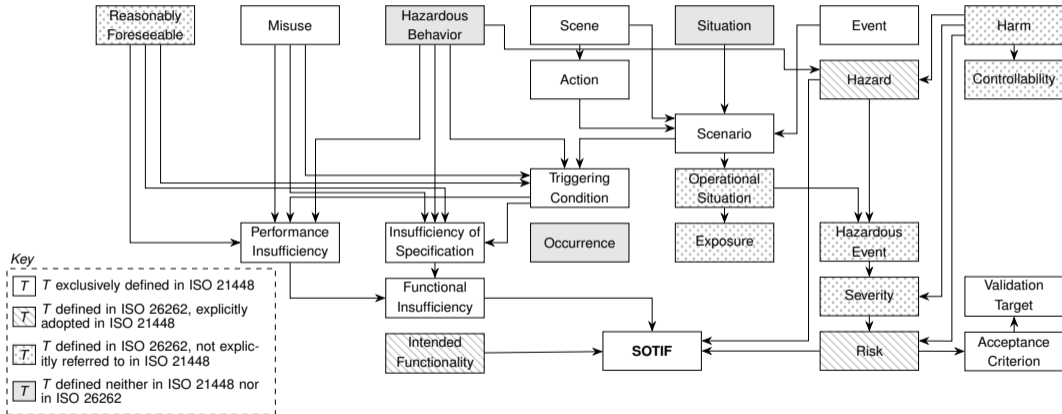
- (1) study and adjust the ISO 21448's terminological risk framework
- (2) examine the relevant normative and informative parts on SOTIF validation and provide constructive suggestions for improvement

How does the ISO 21448:2022 require or suggest performing a quantitative SOTIF validation?

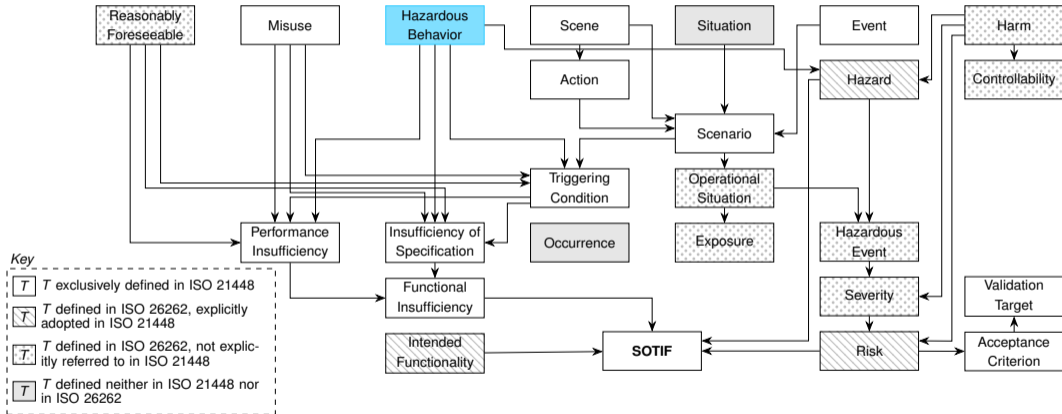
To answer this, we...

- (1) study and adjust the ISO 21448's terminological risk framework
- (2) examine the relevant normative and informative parts on SOTIF validation and provide constructive suggestions for improvement

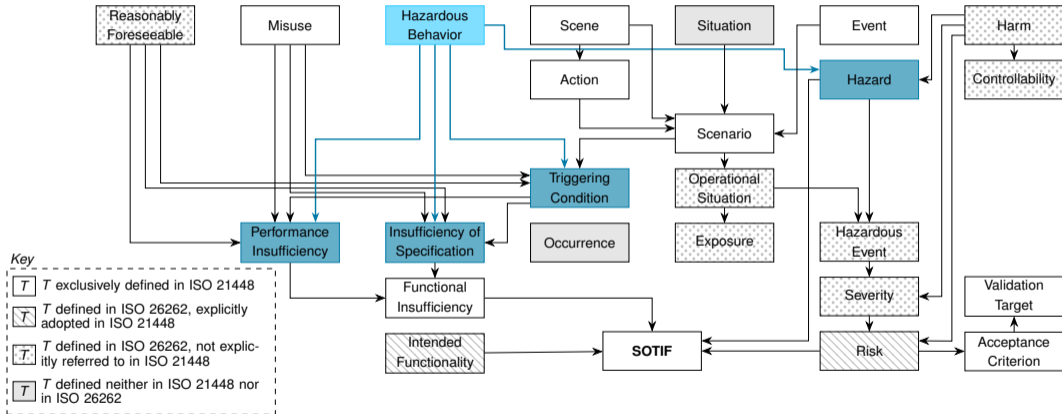
Relations of Definitions



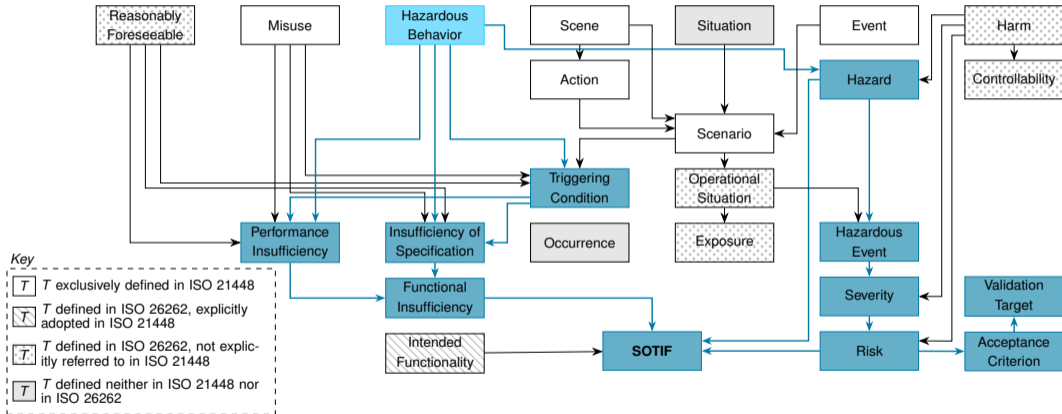
Relations of Definitions



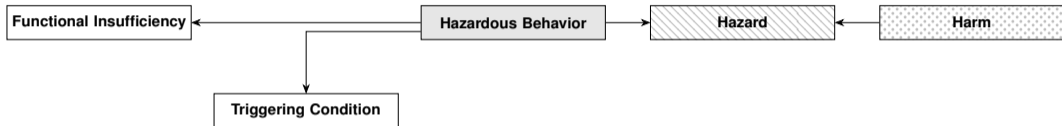
Relations of Definitions



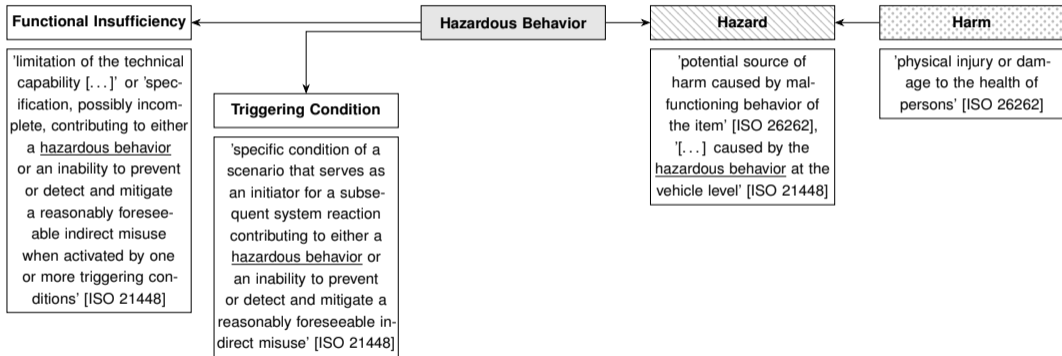
Relations of Definitions



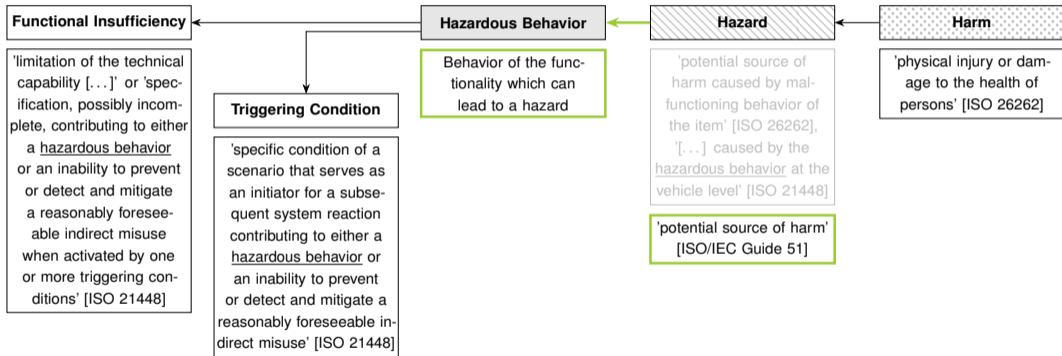
Relations of Definitions



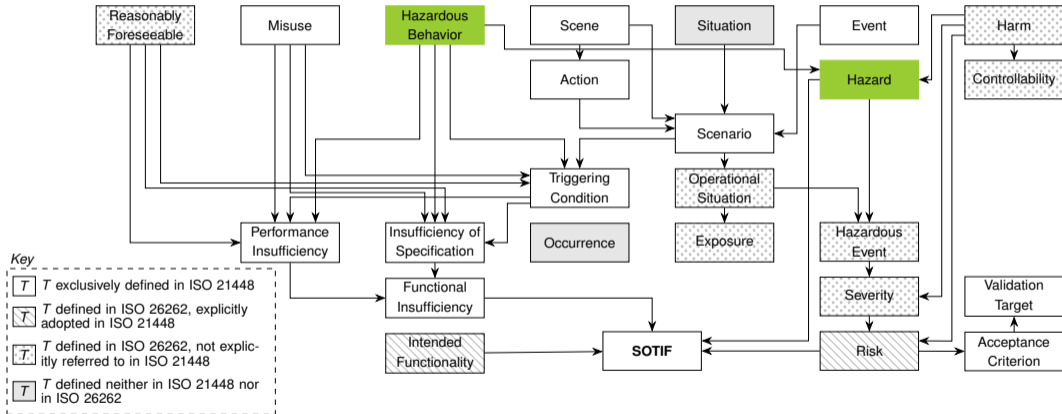
Relations of Definitions



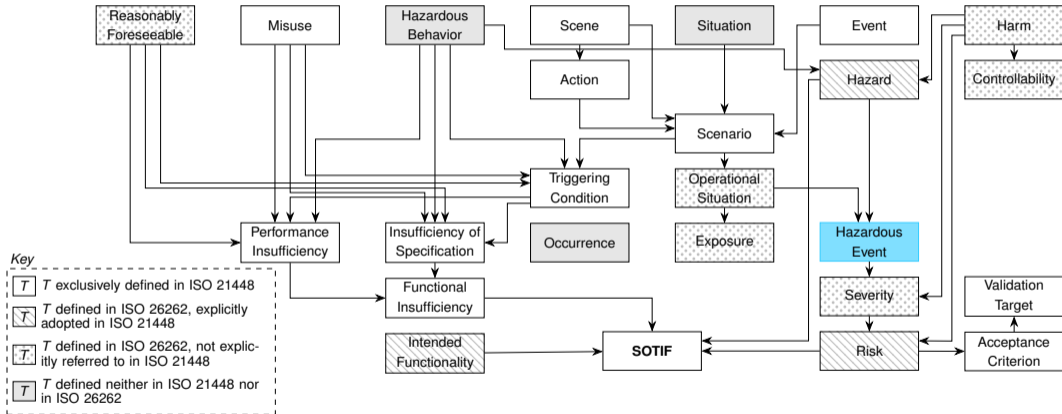
Relations of Definitions



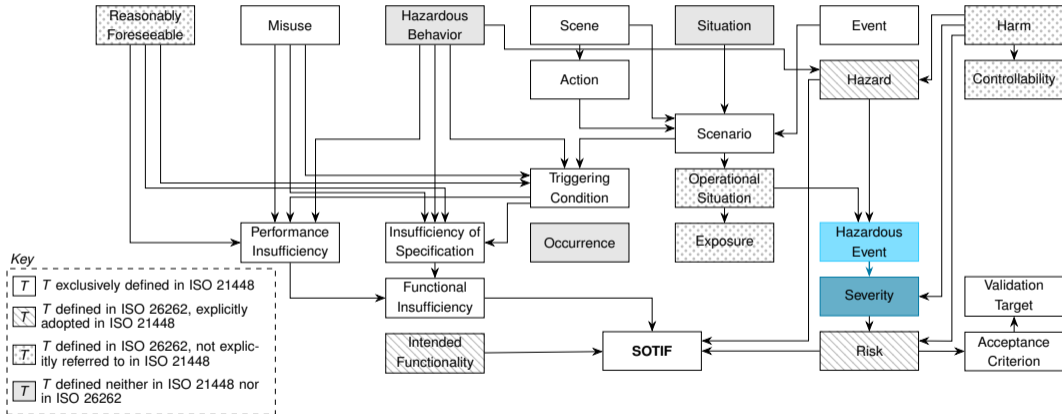
Relations of Definitions



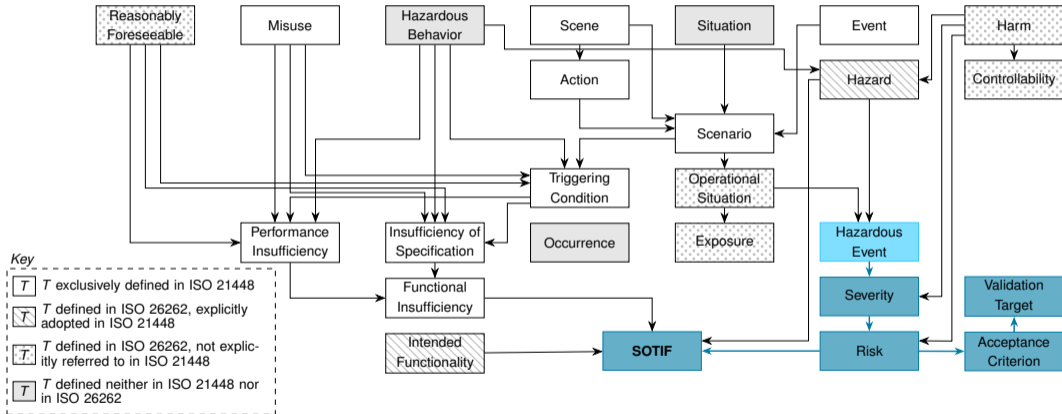
Relations of Definitions



Relations of Definitions



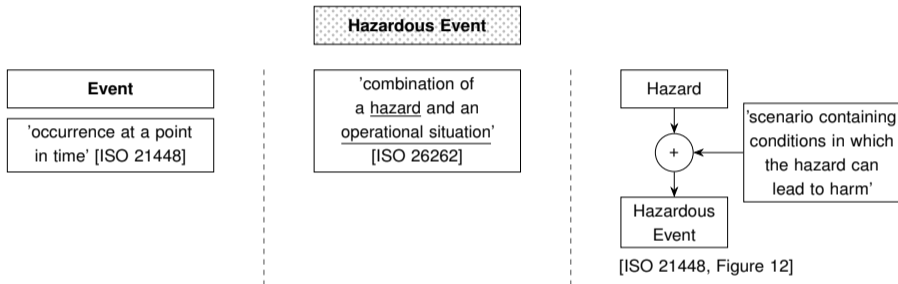
Relations of Definitions



Hazardous Event

'combination of
a hazard and an
operational situation'
[ISO 26262]

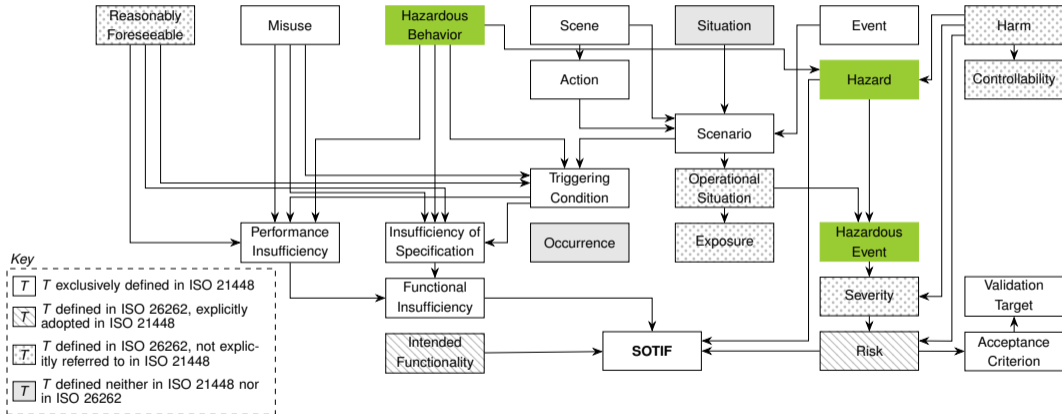
Relations of Definitions



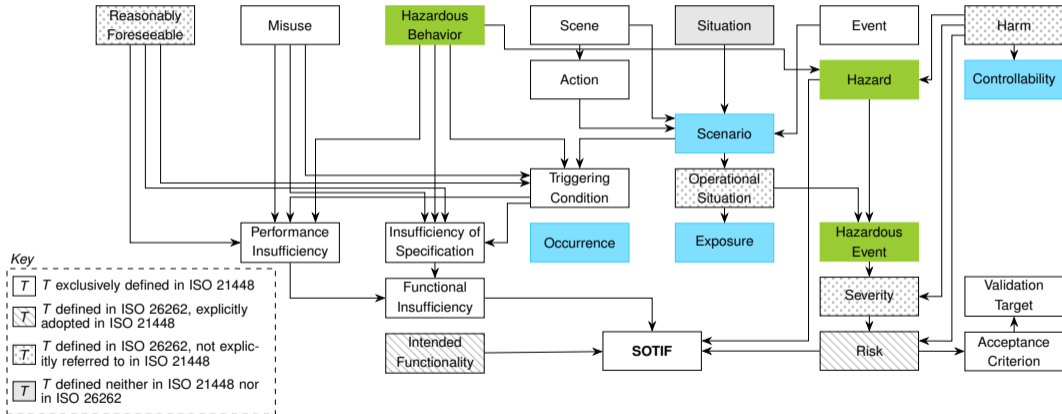


event that is a combination of a hazard and a scenario containing conditions in which the hazard can lead to harm

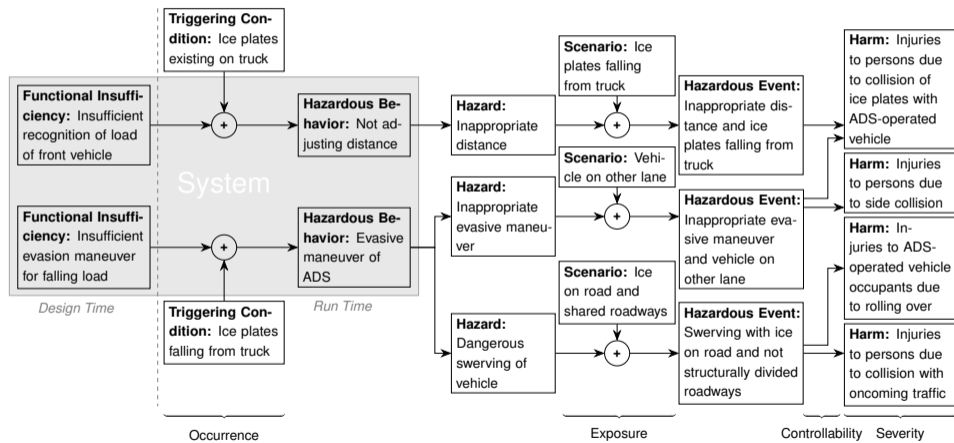
Relations of Definitions



Relations of Definitions



Example of the Terminological Risk Framework



How does the ISO 21448:2022 require or suggest performing a quantitative SOTIF validation?

To answer this, we...

- (1) study and adjust the ISO 21448's terminological risk framework
- (2) examine the relevant normative and informative parts on SOTIF validation and provide constructive suggestions for improvement

How does the ISO 21448:2022 require or suggest performing a quantitative SOTIF validation?

To answer this, we...

- (1) study and adjust the ISO 21448's terminological risk framework
- (2) examine the relevant normative and informative parts on SOTIF validation and provide constructive suggestions for improvement

Relevant clauses within the normative part:

Relevant clauses within the normative part:

Clause 6: Identification and evaluation of hazards

Relevant clauses within the normative part:

Clause 6: Identification and evaluation of hazards

Clause 7: Identification and evaluation of potential functional insufficiencies and potential triggering conditions

Relevant clauses within the normative part:

Clause 6: Identification and evaluation of hazards

Clause 7: Identification and evaluation of potential functional insufficiencies and potential triggering conditions

Clause 9: Definition of the verification and validation strategy

Relevant clauses within the normative part:

Clause 6: Identification and evaluation of hazards

Clause 7: Identification and evaluation of potential functional insufficiencies and potential triggering conditions

Clause 9: Definition of the verification and validation strategy

Remark: The normative part of the ISO 21448 is rather sparse with requirements compared to other standards

Clause 6: Identification and evaluation of hazards

Clause 6: Identification and evaluation of hazards

- no ASIL classification required unlike in the ISO 26262

Clause 6: Identification and evaluation of hazards

- no ASIL classification required unlike in the ISO 26262
- idea of considering severity, exposure and controllability remains

Clause 6: Identification and evaluation of hazards

- no ASIL classification required unlike in the ISO 26262
- idea of considering severity, exposure and controllability remains
 - severity, controllability: treated as binary variables

Clause 6: Identification and evaluation of hazards

- no ASIL classification required unlike in the ISO 26262
- idea of considering severity, exposure and controllability remains
 - severity, controllability: treated as binary variables
 - exposure: not considered

Clause 6: Identification and evaluation of hazards

- no ASIL classification required unlike in the ISO 26262
- idea of considering severity, exposure and controllability remains
 - severity, controllability: treated as binary variables
 - exposure: not considered
- ✗ general reasoning why this simplification of the ASIL classification should be sufficient is missing

Clause 6: Identification and evaluation of hazards

- no ASIL classification required unlike in the ISO 26262
- idea of considering severity, exposure and controllability remains
 - severity, controllability: treated as binary variables
 - exposure: not considered
 - ✗ general reasoning why this simplification of the ASIL classification should be sufficient is missing
- acceptance criteria must be formulated for SOTIF-related hazardous events

Clause 6: Identification and evaluation of hazards

- no ASIL classification required unlike in the ISO 26262
- idea of considering severity, exposure and controllability remains
 - severity, controllability: treated as binary variables
 - exposure: not considered
 - ✗ general reasoning why this simplification of the ASIL classification should be sufficient is missing
- acceptance criteria must be formulated for SOTIF-related hazardous events
 - both qualitative and quantitative acceptance criteria are permitted

Clause 6: Identification and evaluation of hazards

- no ASIL classification required unlike in the ISO 26262
- idea of considering severity, exposure and controllability remains
 - severity, controllability: treated as binary variables
 - exposure: not considered
 - ✗ general reasoning why this simplification of the ASIL classification should be sufficient is missing
- acceptance criteria must be formulated for SOTIF-related hazardous events
 - both qualitative and quantitative acceptance criteria are permitted
 - quantitative acceptance criteria are exclusively mentioned: GAMAB, PRB, ALARP, MEM

Clause 7: Identification and evaluation of potential functional insufficiencies and potential triggering conditions

Clause 7: Identification and evaluation of potential functional insufficiencies and potential triggering conditions

- systematic qualitative or quantitative analysis of potential functional insufficiencies and associated triggering conditions demanded

Clause 7: Identification and evaluation of potential functional insufficiencies and potential triggering conditions

- systematic qualitative or quantitative analysis of potential functional insufficiencies and associated triggering conditions demanded
- for scenarios containing identified triggering conditions SOTIF-achievability needs to be demonstrated

Clause 9: Definition of the verification and validation strategy

Clause 9: Definition of the verification and validation strategy

- validation targets should be derived to argue that the acceptance criteria are fulfilled

Clause 9: Definition of the verification and validation strategy

- validation targets should be derived to argue that the acceptance criteria are fulfilled
- strategy to provide evidence that validation targets are met must be provided

Clause 9: Definition of the verification and validation strategy

- validation targets should be derived to argue that the acceptance criteria are fulfilled
- strategy to provide evidence that validation targets are met must be provided
- an example for deriving a validation target from a given quantitative acceptance criterion is given in Annex C.2 leading to the following decomposition

$$A_H = R_{HB} \cdot P_{E|HB} \cdot P_{C|E} \cdot P_{S|C}$$

Clause 9: Definition of the verification and validation strategy

- validation targets should be derived to argue that the acceptance criteria are fulfilled
- strategy to provide evidence that validation targets are met must be provided
- an example for deriving a validation target from a given quantitative acceptance criterion is given in Annex C.2 leading to the following decomposition

$$A_H = R_{HB} \cdot P_{E|HB} \cdot P_{C|E} \cdot P_{S|C}$$

✗ deficient use of conditional probabilities

Clause 9: Definition of the verification and validation strategy

- validation targets should be derived to argue that the acceptance criteria are fulfilled
- strategy to provide evidence that validation targets are met must be provided
- an example for deriving a validation target from a given quantitative acceptance criterion is given in Annex C.2 leading to the following decomposition

$$A_H = R_{HB} \cdot P_{E|HB} \cdot P_{C|E} \cdot P_{S|C}$$

- ✗ deficient use of conditional probabilities
- ✗ probabilities are claimed to be known from field data

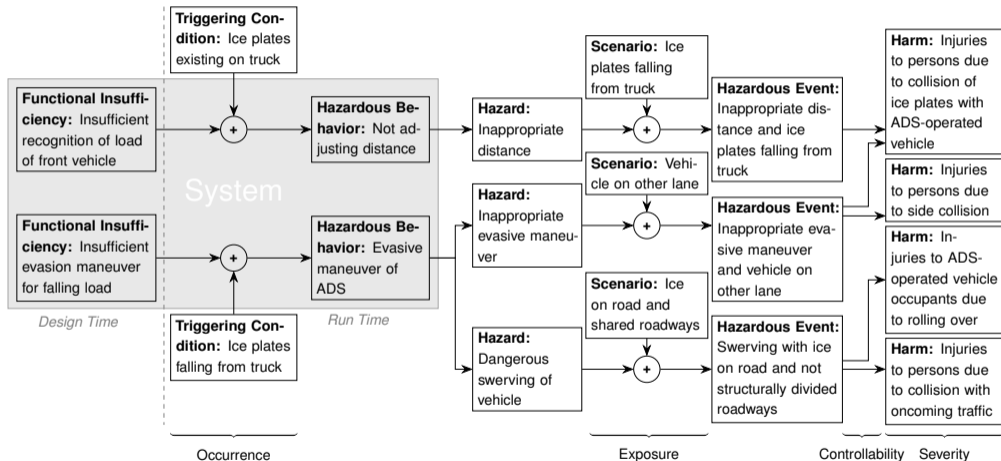
Clause 9: Definition of the verification and validation strategy

- validation targets should be derived to argue that the acceptance criteria are fulfilled
- strategy to provide evidence that validation targets are met must be provided
- an example for deriving a validation target from a given quantitative acceptance criterion is given in Annex C.2 leading to the following decomposition

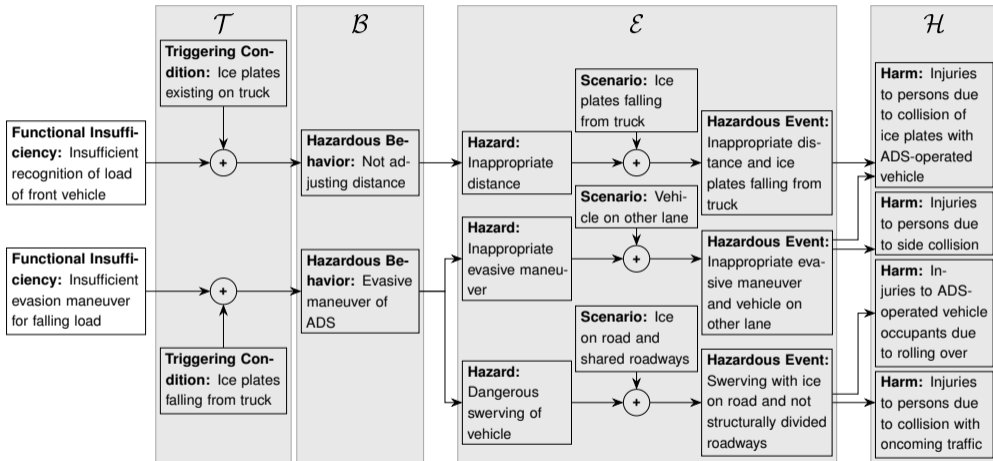
$$A_H = R_{HB} \cdot P_{E|HB} \cdot P_{C|E} \cdot P_{S|C}$$

- ✗ deficient use of conditional probabilities
- ✗ probabilities are claimed to be known from field data
- ✗ 1-to-1 relation between hazardous behavior and harm is implicitly assumed

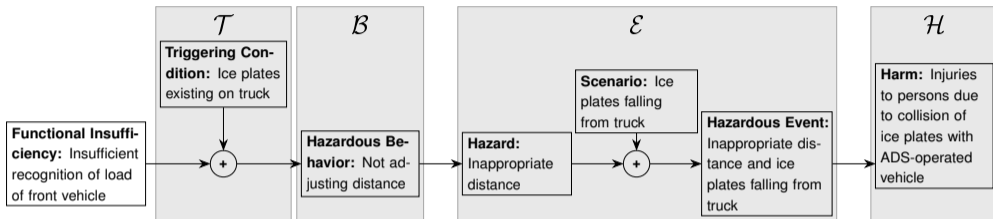
Validation of the SOTIF using Quantitative Acceptance Criteria



Validation of the SOTIF using Quantitative Acceptance Criteria



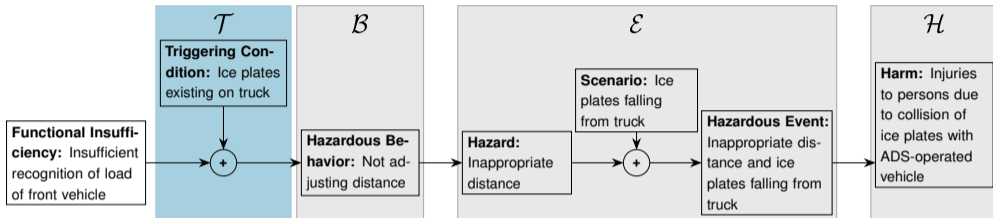
Validation of the SOTIF using Quantitative Acceptance Criteria



- Probability of occurrence of a given harm \mathcal{H} :

$$P(\mathcal{H}) \leq \sum_{\mathcal{E}, \mathcal{B}, \mathcal{T}} P(\mathcal{T})P(\mathcal{B}|\mathcal{T})P(\mathcal{E}|\mathcal{B}, \mathcal{T})P(\mathcal{H}|\mathcal{E}, \mathcal{B}, \mathcal{T})$$

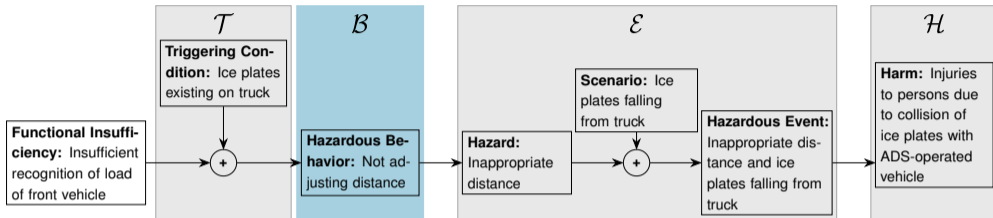
Validation of the SOTIF using Quantitative Acceptance Criteria



- Probability of occurrence of a given harm \mathcal{H} :

$$P(\mathcal{H}) \leq \sum_{\mathcal{E}, \mathcal{B}, \mathcal{T}} P(\mathcal{T}) P(\mathcal{B}|\mathcal{T}) P(\mathcal{E}|\mathcal{B}, \mathcal{T}) P(\mathcal{H}|\mathcal{E}, \mathcal{B}, \mathcal{T})$$

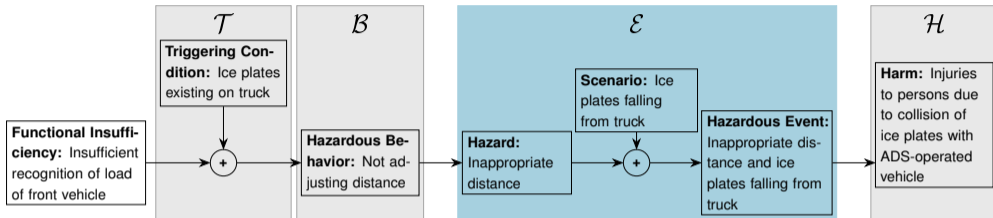
Validation of the SOTIF using Quantitative Acceptance Criteria



- Probability of occurrence of a given harm \mathcal{H} :

$$P(\mathcal{H}) \leq \sum_{\mathcal{E}, \mathcal{B}, \mathcal{T}} P(\mathcal{T})P(\mathcal{B}|\mathcal{T})P(\mathcal{E}|\mathcal{B}, \mathcal{T})P(\mathcal{H}|\mathcal{E}, \mathcal{B}, \mathcal{T})$$

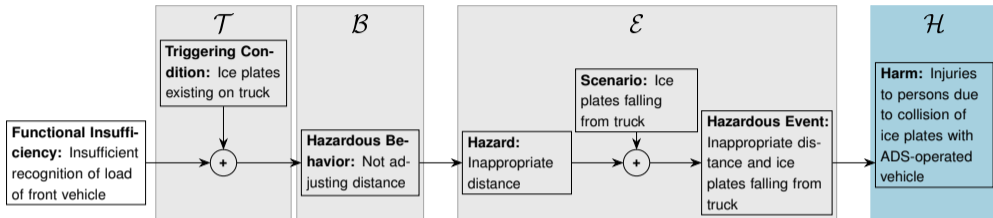
Validation of the SOTIF using Quantitative Acceptance Criteria



- Probability of occurrence of a given harm \mathcal{H} :

$$P(\mathcal{H}) \leq \sum_{\mathcal{E}, \mathcal{B}, \mathcal{T}} P(\mathcal{T})P(\mathcal{B}|\mathcal{T})P(\mathcal{E}|\mathcal{B}, \mathcal{T})P(\mathcal{H}|\mathcal{E}, \mathcal{B}, \mathcal{T})$$

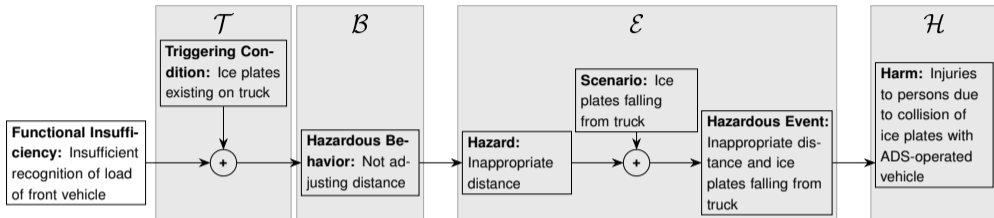
Validation of the SOTIF using Quantitative Acceptance Criteria



- Probability of occurrence of a given harm \mathcal{H} :

$$P(\mathcal{H}) \leq \sum_{\mathcal{E}, \mathcal{B}, \mathcal{T}} P(\mathcal{T})P(\mathcal{B}|\mathcal{T})P(\mathcal{E}|\mathcal{B}, \mathcal{T})P(\mathcal{H}|\mathcal{E}, \mathcal{B}, \mathcal{T})$$

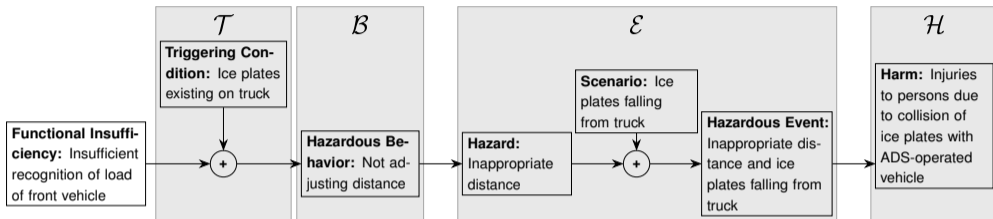
Validation of the SOTIF using Quantitative Acceptance Criteria



- Probability of occurrence of a given harm \mathcal{H} :

$$P(\mathcal{H}) \leq \sum_{\mathcal{E}, \mathcal{B}, \mathcal{T}} P(\mathcal{T})P(\mathcal{B}|\mathcal{T})P(\mathcal{E}|\mathcal{B}, \mathcal{T})P(\mathcal{H}|\mathcal{E}, \mathcal{B}, \mathcal{T})$$

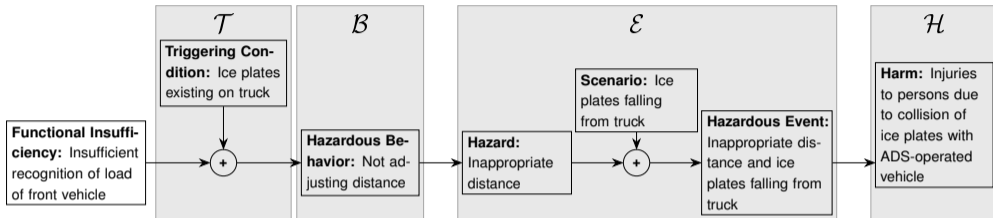
Validation of the SOTIF using Quantitative Acceptance Criteria



- Probability of occurrence of a given harm \mathcal{H} :

$$P(\mathcal{H}) \leq \sum_{\mathcal{E}, \mathcal{B}, \mathcal{T}} P(\mathcal{T})P(\mathcal{B}|\mathcal{T})P(\mathcal{E}|\mathcal{B}, \mathcal{T})P(\mathcal{H}|\mathcal{E}, \mathcal{B}, \mathcal{T})$$

Validation of the SOTIF using Quantitative Acceptance Criteria



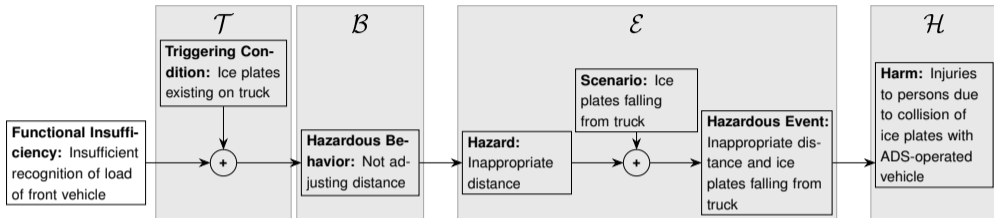
- Probability of occurrence of a given harm \mathcal{H} :

$$P(\mathcal{H}) \leq \sum_{\mathcal{E}, \mathcal{B}, \mathcal{T}} P(\mathcal{T})P(\mathcal{B}|\mathcal{T})P(\mathcal{E}|\mathcal{B}, \mathcal{T})P(\mathcal{H}|\mathcal{E}, \mathcal{B}, \mathcal{T})$$

- Probability of occurrence of a given harm \mathcal{H} in combination with a severity level \mathcal{S} :

$$P(\mathcal{H}, \mathcal{S}) \leq \sum_{\mathcal{E}, \mathcal{B}, \mathcal{T}} P(\mathcal{T})P(\mathcal{B}|\mathcal{T})P(\mathcal{E}|\mathcal{B}, \mathcal{T})P(\mathcal{H}|\mathcal{E}, \mathcal{B}, \mathcal{T})P(\mathcal{S}|\mathcal{H}, \mathcal{E}, \mathcal{B}, \mathcal{T})$$

Validation of the SOTIF using Quantitative Acceptance Criteria



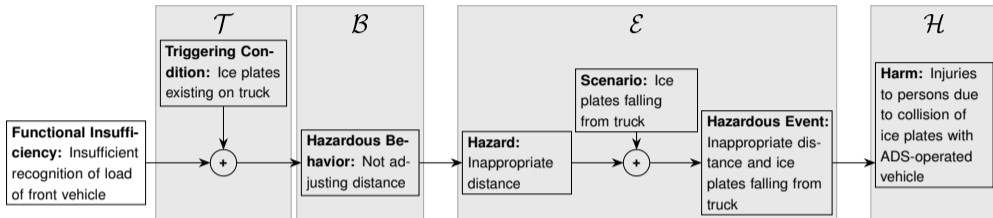
- Probability of occurrence of a given harm \mathcal{H} :

$$P(\mathcal{H}) \leq \sum_{\mathcal{E}, \mathcal{B}, \mathcal{T}} P(\mathcal{T})P(\mathcal{B}|\mathcal{T})P(\mathcal{E}|\mathcal{B}, \mathcal{T})P(\mathcal{H}|\mathcal{E}, \mathcal{B}, \mathcal{T})$$

- Probability of occurrence of a given harm \mathcal{H} in combination with a severity level \mathcal{S} :

$$P(\mathcal{H}, \mathcal{S}) \leq \sum_{\mathcal{E}, \mathcal{B}, \mathcal{T}} P(\mathcal{T})P(\mathcal{B}|\mathcal{T})P(\mathcal{E}|\mathcal{B}, \mathcal{T})P(\mathcal{H}|\mathcal{E}, \mathcal{B}, \mathcal{T})P(\mathcal{S}|\mathcal{H}, \mathcal{E}, \mathcal{B}, \mathcal{T})$$

Validation of the SOTIF using Quantitative Acceptance Criteria



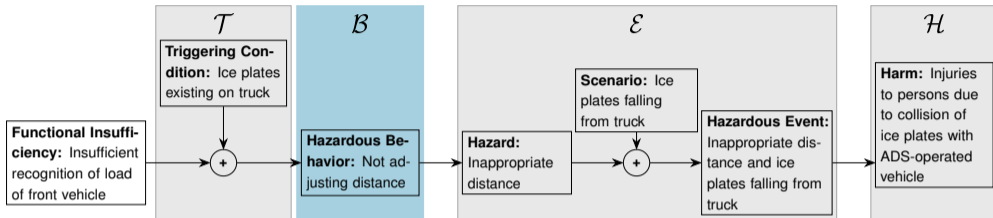
✗ Decomposition given in the Annex C.2 of the ISO 21448:

$$A_H = R_{HB} \cdot P_{E|HB} \cdot P_{C|E} \cdot P_{S|C}$$

✓ Approach proposed:

$$P(\mathcal{H}, S) \leq \sum_{\mathcal{E}, \mathcal{B}, \mathcal{T}} P(\mathcal{T}) P(\mathcal{B}|\mathcal{T}) P(\mathcal{E}|\mathcal{B}, \mathcal{T}) P(\mathcal{H}|\mathcal{E}, \mathcal{B}, \mathcal{T}) P(S|\mathcal{H}, \mathcal{E}, \mathcal{B}, \mathcal{T})$$

Validation of the SOTIF using Quantitative Acceptance Criteria



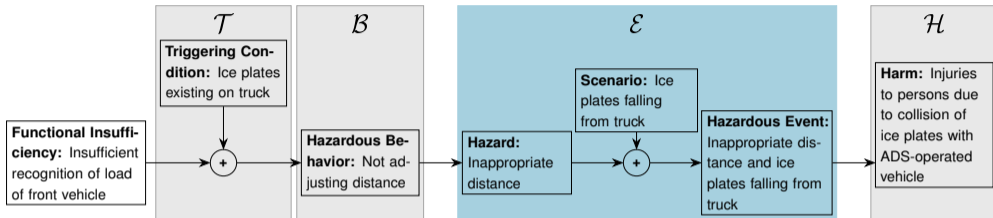
✗ Decomposition given in the Annex C.2 of the ISO 21448:

$$A_H = R_{HB} \cdot P_{E|HB} \cdot P_{C|E} \cdot P_{S|C}$$

✓ Approach proposed:

$$P(\mathcal{H}, S) \leq \sum_{\mathcal{E}, \mathcal{B}, \mathcal{T}} P(\mathcal{T}) P(\mathcal{B}|\mathcal{T}) P(\mathcal{E}|\mathcal{B}, \mathcal{T}) P(\mathcal{H}|\mathcal{E}, \mathcal{B}, \mathcal{T}) P(S|\mathcal{H}, \mathcal{E}, \mathcal{B}, \mathcal{T})$$

Validation of the SOTIF using Quantitative Acceptance Criteria



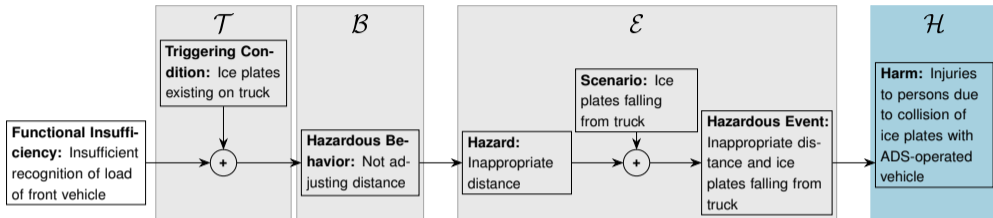
✗ Decomposition given in the Annex C.2 of the ISO 21448:

$$A_H = R_{HB} \cdot P_{E|HB} \cdot P_{C|E} \cdot P_{S|C}$$

✓ Approach proposed:

$$P(\mathcal{H}, S) \leq \sum_{\mathcal{E}, \mathcal{B}, \mathcal{T}} P(\mathcal{T}) P(\mathcal{B}|\mathcal{T}) P(\mathcal{E}|\mathcal{B}, \mathcal{T}) P(\mathcal{H}|\mathcal{E}, \mathcal{B}, \mathcal{T}) P(S|\mathcal{H}, \mathcal{E}, \mathcal{B}, \mathcal{T})$$

Validation of the SOTIF using Quantitative Acceptance Criteria



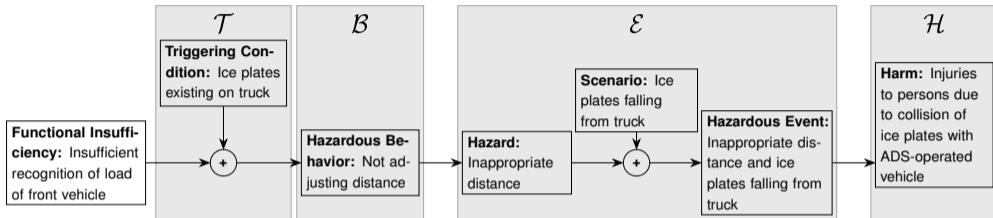
✗ Decomposition given in the Annex C.2 of the ISO 21448:

$$A_H = R_{HB} \cdot P_{E|HB} \cdot P_{C|E} \cdot P_{S|C}$$

✓ Approach proposed:

$$P(\mathcal{H}, S) \leq \sum_{\mathcal{E}, \mathcal{B}, \mathcal{T}} P(\mathcal{T}) P(\mathcal{B}|\mathcal{T}) P(\mathcal{E}|\mathcal{B}, \mathcal{T}) P(\mathcal{H}|\mathcal{E}, \mathcal{B}, \mathcal{T}) P(S|\mathcal{H}, \mathcal{E}, \mathcal{B}, \mathcal{T})$$

Validation of the SOTIF using Quantitative Acceptance Criteria



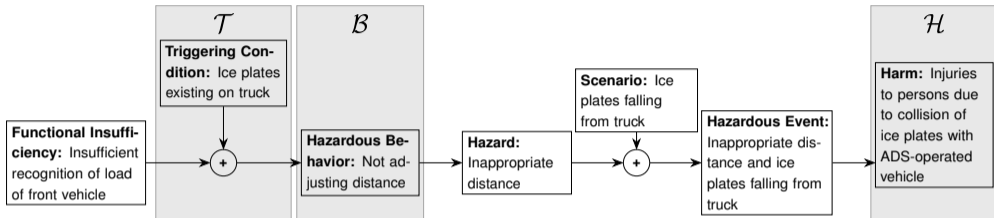
✗ Decomposition given in the Annex C.2 of the ISO 21448:

$$A_H = R_{HB} \cdot P_{E|HB} \cdot P_{C|E} \cdot P_{S|C}$$

✓ Approach proposed:

$$P(\mathcal{H}, S) \leq \sum_{\mathcal{E}, \mathcal{B}, \mathcal{T}} P(\mathcal{T})P(\mathcal{B}|\mathcal{T})P(\mathcal{E}|\mathcal{B}, \mathcal{T})P(\mathcal{H}|\mathcal{E}, \mathcal{B}, \mathcal{T})P(S|\mathcal{H}, \mathcal{E}, \mathcal{B}, \mathcal{T})$$

Validation of the SOTIF using Quantitative Acceptance Criteria

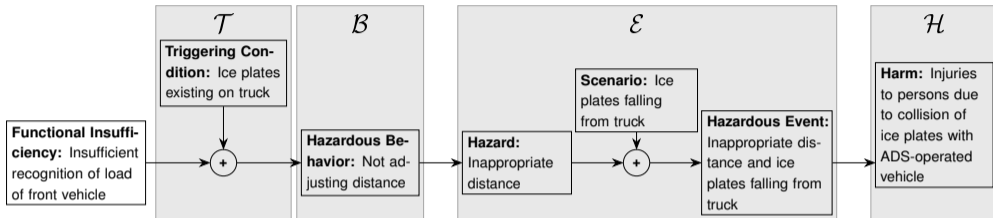


- Other discretizations are also conceivable, for example:

$$P(\mathcal{H}) \leq \sum_{\mathcal{B}, \mathcal{T}} P(\mathcal{T})P(\mathcal{B}|\mathcal{T})P(\mathcal{H}|\mathcal{B}, \mathcal{T})$$

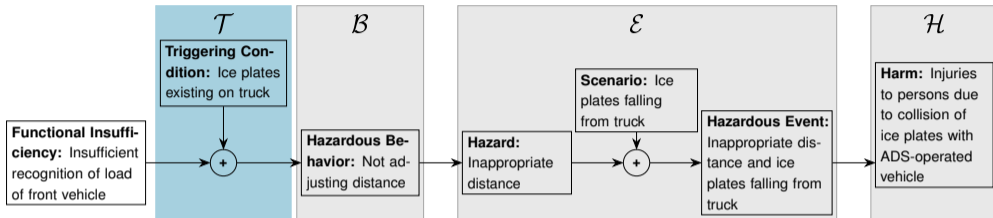
$$P(\mathcal{H}, \mathcal{S}) \leq \sum_{\mathcal{B}, \mathcal{T}} P(\mathcal{T})P(\mathcal{B}|\mathcal{T})P(\mathcal{H}|\mathcal{B}, \mathcal{T})P(\mathcal{S}|\mathcal{H}, \mathcal{B}, \mathcal{T})$$

Validation of the SOTIF using Quantitative Acceptance Criteria



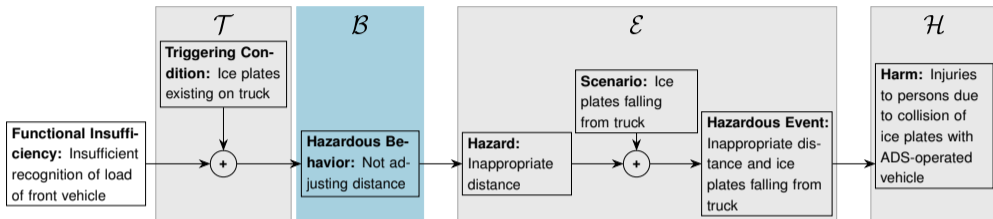
	$P(\mathcal{T})$	$P(\mathcal{B} \mathcal{T})$	$P(\mathcal{E} \mathcal{B}, \mathcal{T})$	$P(\mathcal{H} \mathcal{E}, \mathcal{B}, \mathcal{T})$	$P(\mathcal{S} \mathcal{H}, \mathcal{E}, \mathcal{B}, \mathcal{T})$
Traffic Data					
Proving Ground					
Simulation					

Validation of the SOTIF using Quantitative Acceptance Criteria



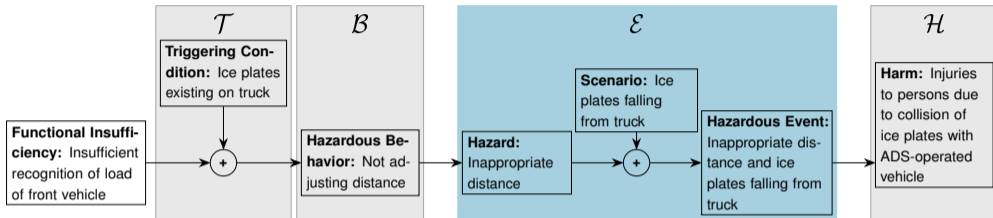
	$P(\mathcal{T})$	$P(\mathcal{B} \mathcal{T})$	$P(\mathcal{E} \mathcal{B}, \mathcal{T})$	$P(\mathcal{H} \mathcal{E}, \mathcal{B}, \mathcal{T})$	$P(\mathcal{S} \mathcal{H}, \mathcal{E}, \mathcal{B}, \mathcal{T})$
Traffic Data	X				
Proving Ground	X				
Simulation	X				

Validation of the SOTIF using Quantitative Acceptance Criteria



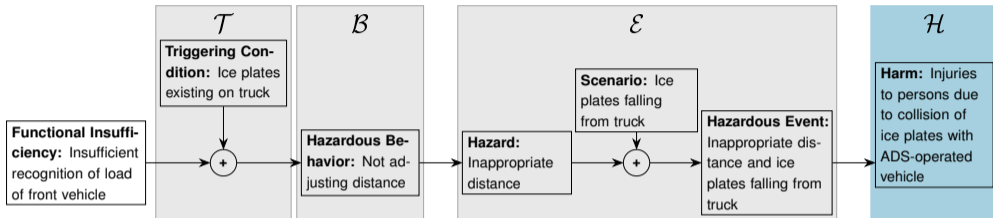
	$P(\mathcal{T})$	$P(\mathcal{B} \mathcal{T})$	$P(\mathcal{E} \mathcal{B}, \mathcal{T})$	$P(\mathcal{H} \mathcal{E}, \mathcal{B}, \mathcal{T})$	$P(\mathcal{S} \mathcal{H}, \mathcal{E}, \mathcal{B}, \mathcal{T})$
Traffic Data	X	X			
Proving Ground	X	✓			
Simulation	X	✓			

Validation of the SOTIF using Quantitative Acceptance Criteria



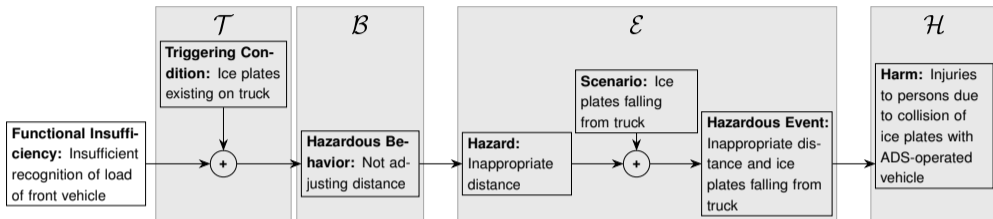
	$P(\mathcal{T})$	$P(\mathcal{B} \mathcal{T})$	$P(\mathcal{E} \mathcal{B}, \mathcal{T})$	$P(\mathcal{H} \mathcal{E}, \mathcal{B}, \mathcal{T})$	$P(\mathcal{S} \mathcal{H}, \mathcal{E}, \mathcal{B}, \mathcal{T})$
Traffic Data	X	X	✓		
Proving Ground	X	✓	(X)		
Simulation	X	✓	(X)		

Validation of the SOTIF using Quantitative Acceptance Criteria



	$P(T)$	$P(B T)$	$P(E B, T)$	$P(H E, B, T)$	$P(S H, E, B, T)$
Traffic Data	X	X	✓	(✓)	
Proving Ground	X	✓	(X)	✓	
Simulation	X	✓	(X)	✓	

Validation of the SOTIF using Quantitative Acceptance Criteria



	$P(\mathcal{T})$	$P(\mathcal{B} \mathcal{T})$	$P(\mathcal{E} \mathcal{B}, \mathcal{T})$	$P(\mathcal{H} \mathcal{E}, \mathcal{B}, \mathcal{T})$	$P(\mathcal{S} \mathcal{H}, \mathcal{E}, \mathcal{B}, \mathcal{T})$
Traffic Data	X	X	✓	(✓)	(✓)
Proving Ground	X	✓	(X)	✓	✓
Simulation	X	✓	(X)	✓	✓

Discussion



- Are there some general rules to derive a suitable decomposition of the risk?

- Are there some general rules to derive a suitable decomposition of the risk?
- Does a scenario-based approach (sufficiently) reduce the validation effort?

- Are there some general rules to derive a suitable decomposition of the risk?
- Does a scenario-based approach (sufficiently) reduce the validation effort?
- Is a quantitative risk assessment possible before employment?

- Are there some general rules to derive a suitable decomposition of the risk?
- Does a scenario-based approach (sufficiently) reduce the validation effort?
- Is a quantitative risk assessment possible before employment?
- How to deal with updates – even post employment?

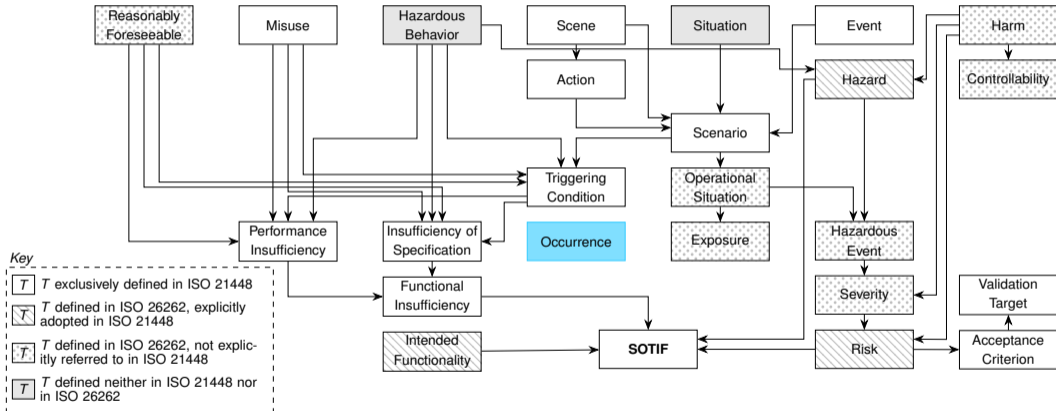
Thank you for the attention.

Contact:

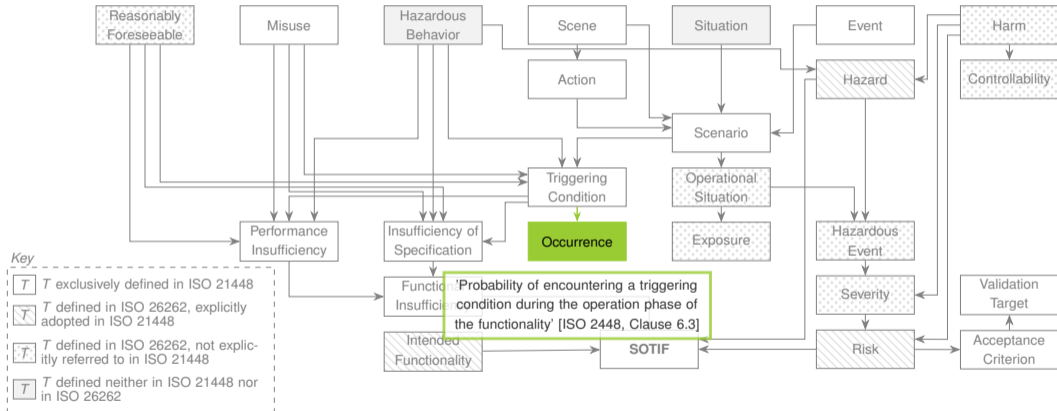
Lina Putze, M.Sc.
German Aerospace Center (DLR) e.V.
Institute of Systems Engineering for Future Mobility
lina.putze@dlr.de



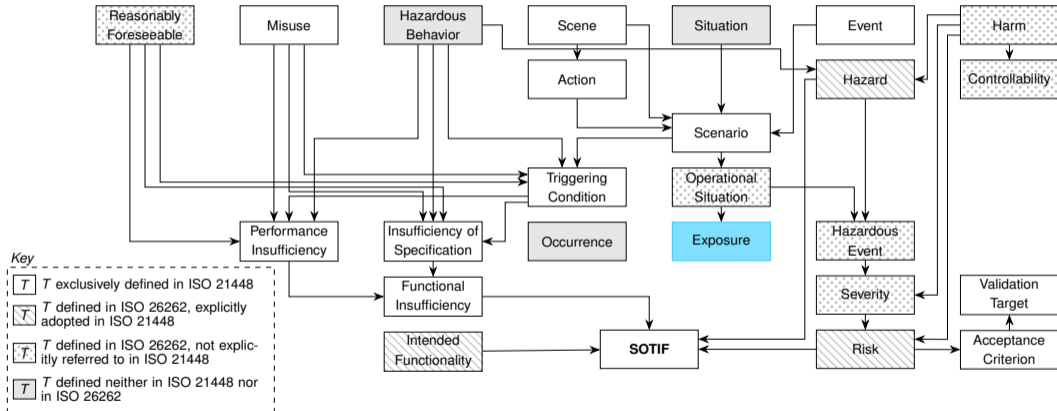
Definition Occurrence



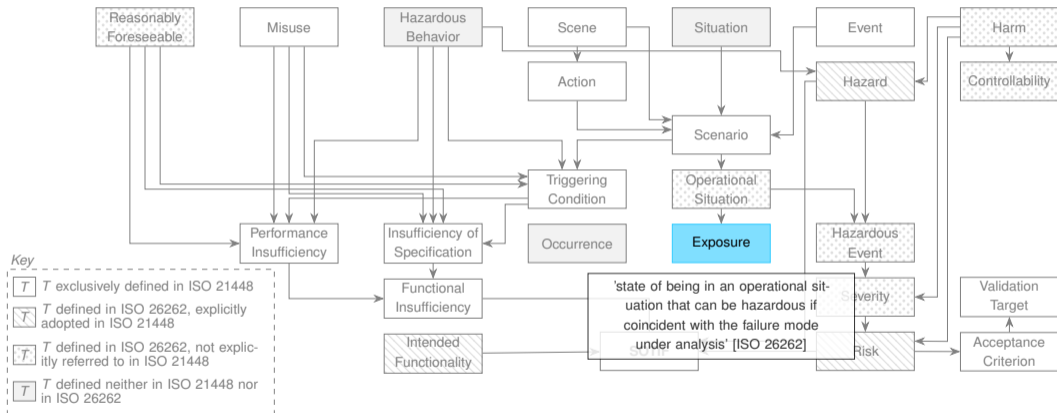
Definition Occurrence



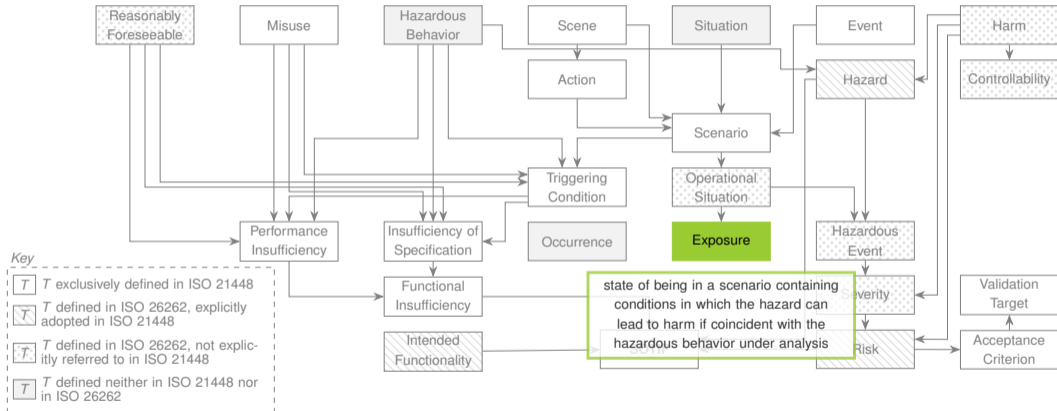
Definition Exposure



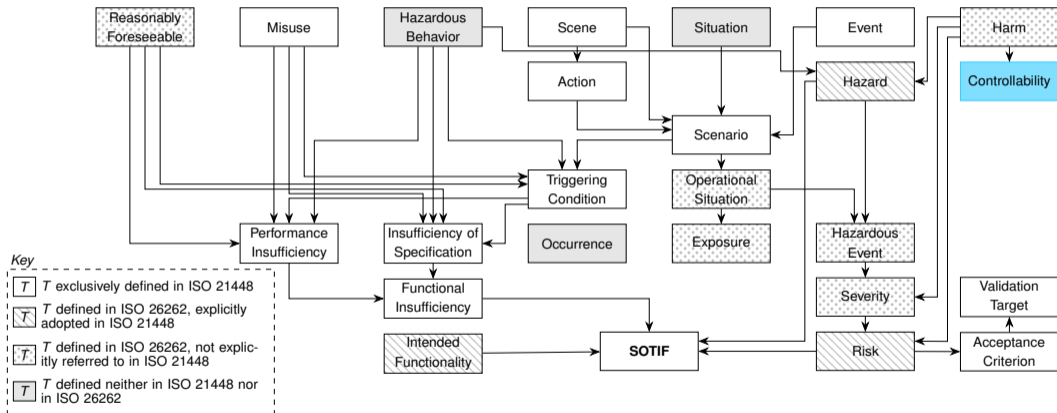
Definition Exposure



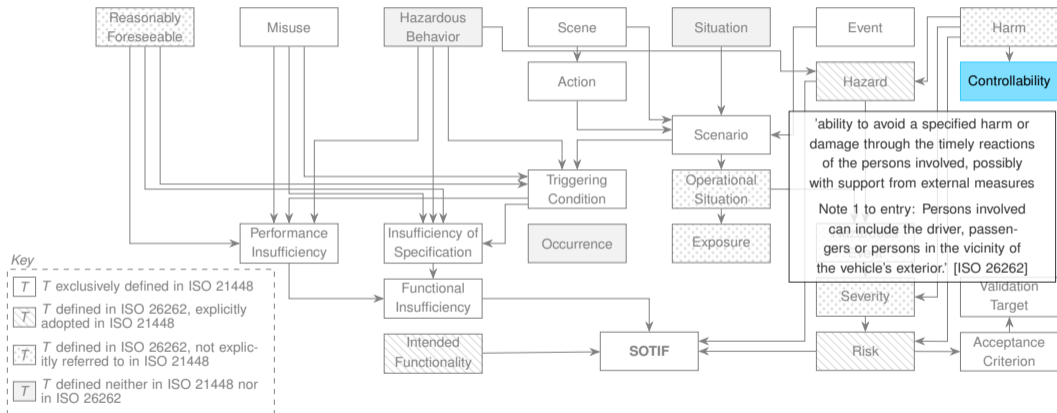
Definition Exposure



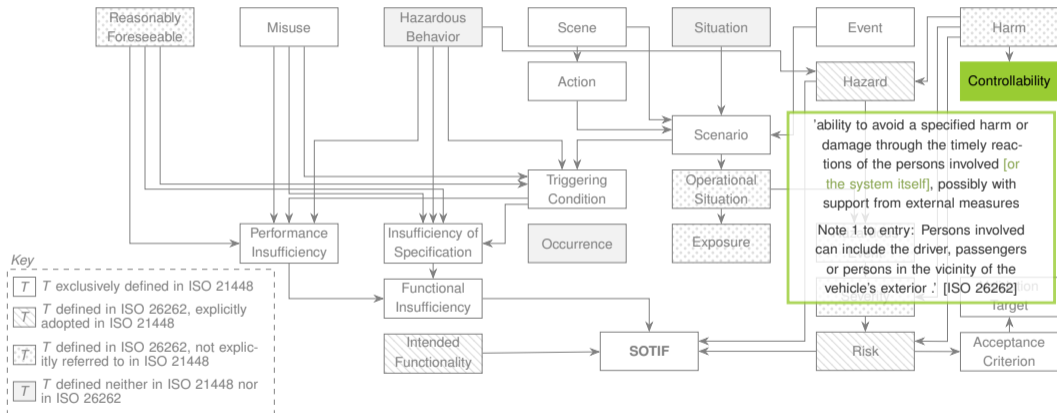
Definition Controllability



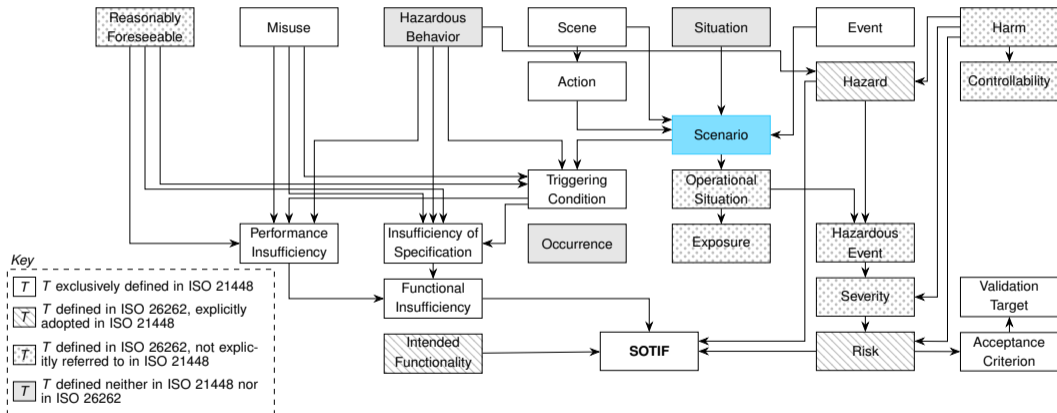
Definition Controllability



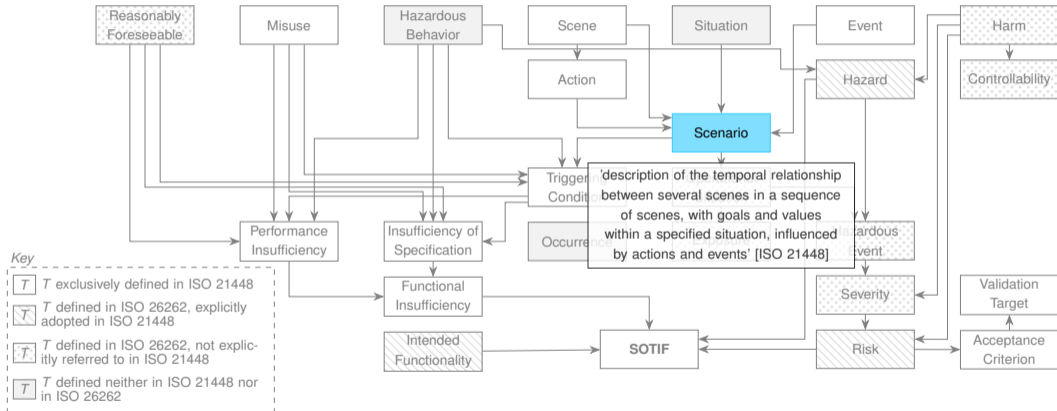
Definition Controllability



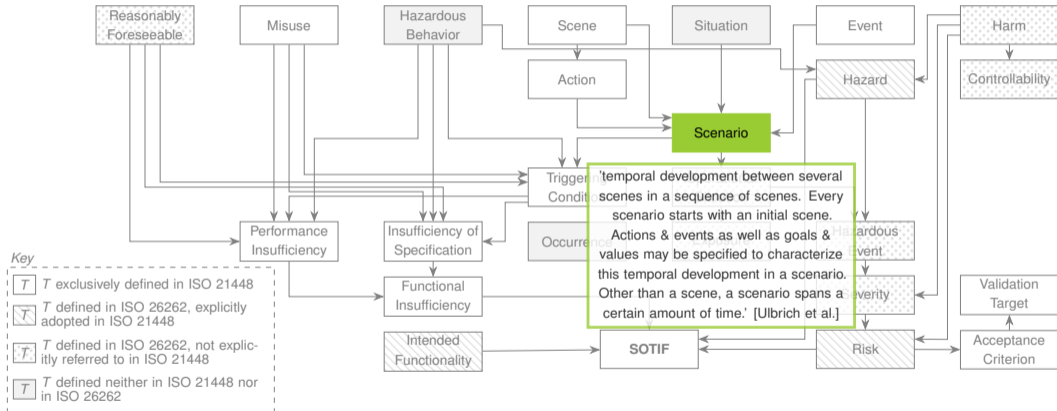
Definition Scenario



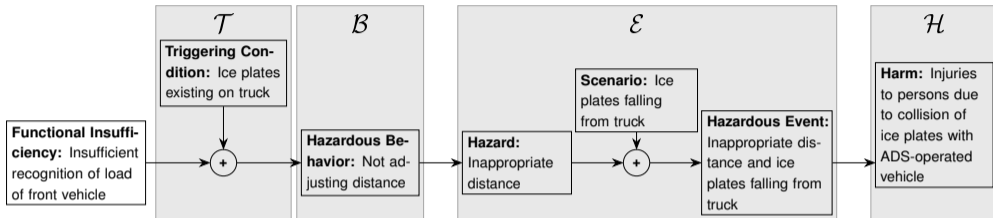
Definition Scenario



Definition Scenario

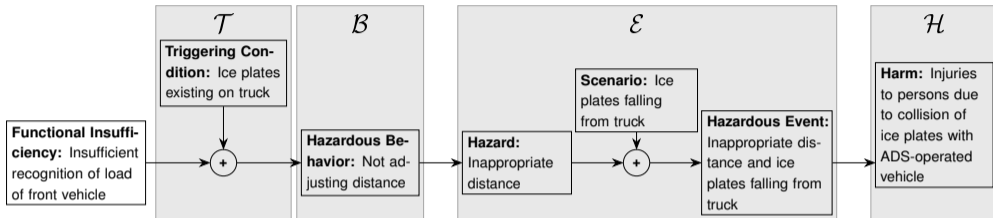


Derivation of Validation Targets



- Suggestion given in the Annex C.2 of the ISO 21448

Derivation of Validation Targets



- Suggestion given in the Annex C.2 of the ISO 21448
 - Solving the factorization of the acceptance criterion A_H for R_{HB} :

$$R_{HB} = \frac{A_H}{P_{E|HB} \cdot P_{C|E} \cdot P_{S|C}}$$

- Estimation of a validation target τ that is sufficient for A_H with confidence level α :

$$\tau = -\ln(1 - \alpha) / R_{HB}$$

- [ISO 21448] International Organization for Standardization, "ISO 21448: Road vehicles – Safety of the intended functionality," 2022.
- [ISO 26262] International Organization for Standardization, "ISO 26262: Road vehicles – Functional safety," 2018.
- [ISO/IEC Guide 51] International Organization for Standardization, "ISO/IEC Guide 51: Safety aspects — Guidelines for their inclusion in standards," 2014.
- [Ulbrich et al.] S. Ulbrich, T. Menzel, A. Reschka, F. Schuldt, and M. Maurer, "Defining and substantiating the terms scene, situation, and scenario for automated driving," in *2015 IEEE 18th international conference on intelligent transportation systems*. IEEE, 2015, pp. 982–988.