Design of a Knowledge-Base Strategy for Capability-Aware Treatment of Uncertainties of Automated Driving Systems

DeJiu Chen, Kenneth Östberg, Matthias Becker, Håkan Sivencrona, Fredrik Warg

1th International Workshop on AI Safety Engineering (**WAISE18**), Safecome18. 18 Sep., Västerås, Sweden.

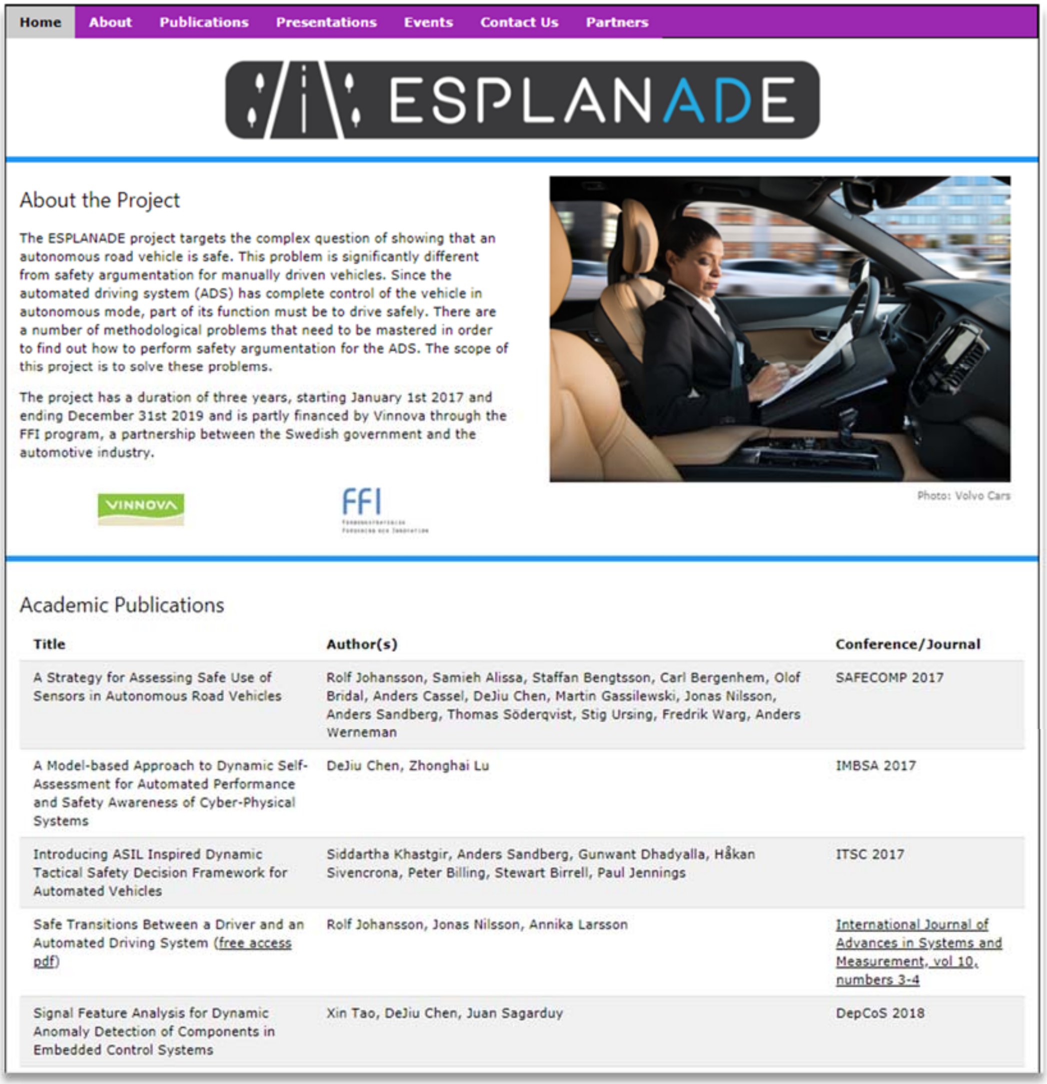
Project ESPLANADE

https://esplanade-project.se/

ESPLANADE

Efficient and Safe
Product Lines of Architectures
Enabling Autonomous Drive

Autoliv · COMENTOR · VOLVO Volvo Group · KTH VETENSKAP OCH KONST · RI. SE · VOLVO · DELPHI · qamcom · semcon · SYSTEMITE



About the Project

The ESPLANADE project targets the complex question of showing that an autonomous road vehicle is safe. This problem is significantly different from safety argumentation for manually driven vehicles. Since the automated driving system (ADS) has complete control of the vehicle in autonomous mode, part of its function must be to drive safely. There are a number of methodological problems that need to be mastered in order to find out how to perform safety argumentation for the ADS. The scope of this project is to solve these problems.

The project has a duration of three years, starting January 1st 2017 and ending December 31st 2019 and is partly financed by Vinnova through the FFI program, a partnership between the Swedish government and the automotive industry.

Photo: Volvo Cars

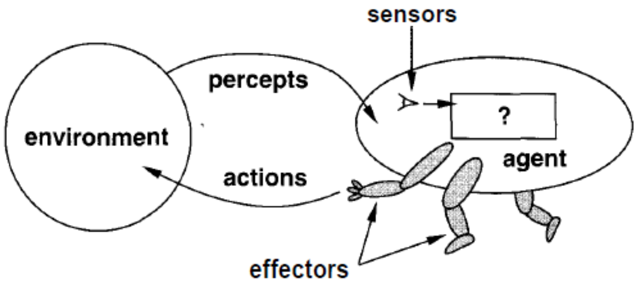
Academic Publications

Title	Author(s)	Conference/Journal
A Strategy for Assessing Safe Use of Sensors in Autonomous Road Vehicles	Rolf Johansson, Samieh Alissa, Staffan Bengtsson, Carl Bergenhem, Olof Bridal, Anders Cassel, DeJiu Chen, Martin Gassilewski, Jonas Nilsson, Anders Sandberg, Thomas Söderqvist, Stig Ursing, Fredrik Warg, Anders Werneman	SAFECOMP 2017
A Model-based Approach to Dynamic Self-Assessment for Automated Performance and Safety Awareness of Cyber-Physical Systems	DeJiu Chen, Zhonghai Lu	IMBSA 2017
Introducing ASIL Inspired Dynamic Tactical Safety Decision Framework for Automated Vehicles	Siddartha Khastgir, Anders Sandberg, Gunwant Dhadyalla, Håkan Sivencrona, Peter Billing, Stewart Birrell, Paul Jennings	ITSC 2017
Safe Transitions Between a Driver and an Automated Driving System (free access pdf)	Rolf Johansson, Jonas Nilsson, Annika Larsson	International Journal of Advances in Systems and Measurement, vol 10, numbers 3-4
Signal Feature Analysis for Dynamic Anomaly Detection of Components in Embedded Control Systems	Xin Tao, DeJiu Chen, Juan Sagarduy	DepCoS 2018

Automated Driving Systems

Intelligent Agents

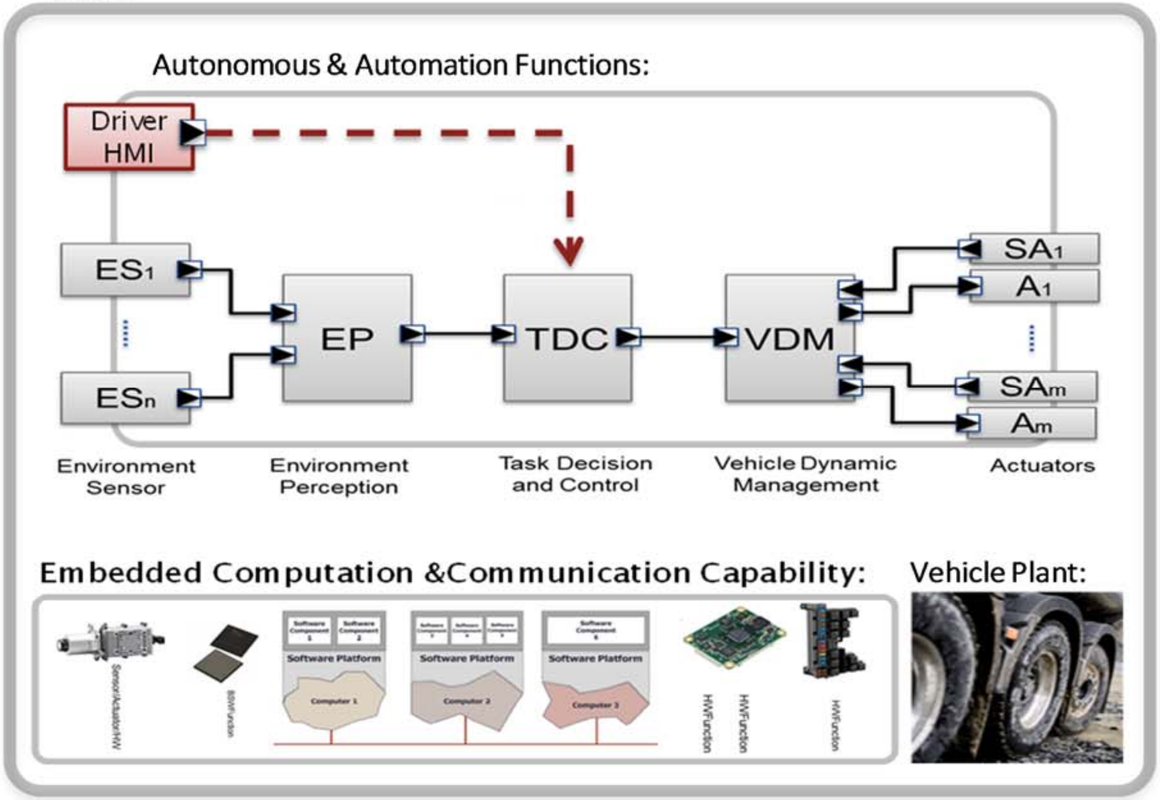
Russell, Stuart J.; Norvig, Peter (1995). Artificial Intelligence: A Modern Approach. Prentice Hall.

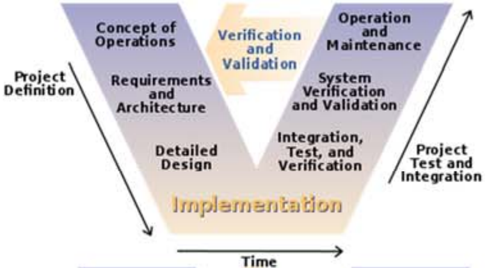


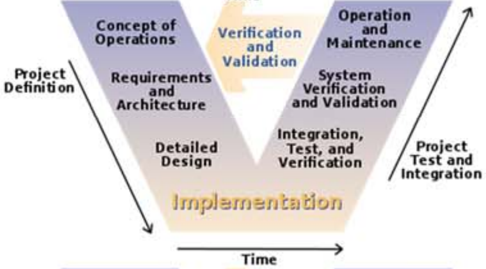
Driver:

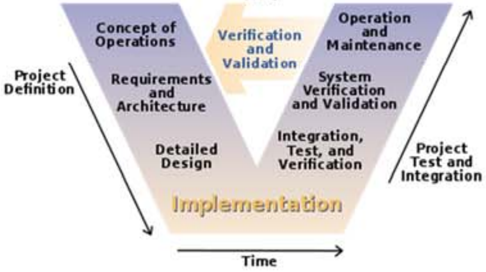
Environment:

Vehicle:



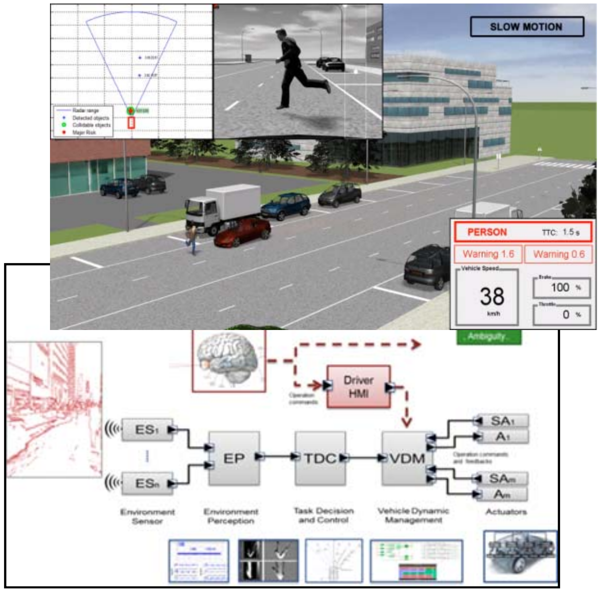

Concept of Operations
Verification and Validation
Operation and Maintenance
Project Definition
Requirements and Architecture
System Verification and Validation
Detailed Design
Integration, Test, and Verification
Project Test and Integration
Implementation
Time

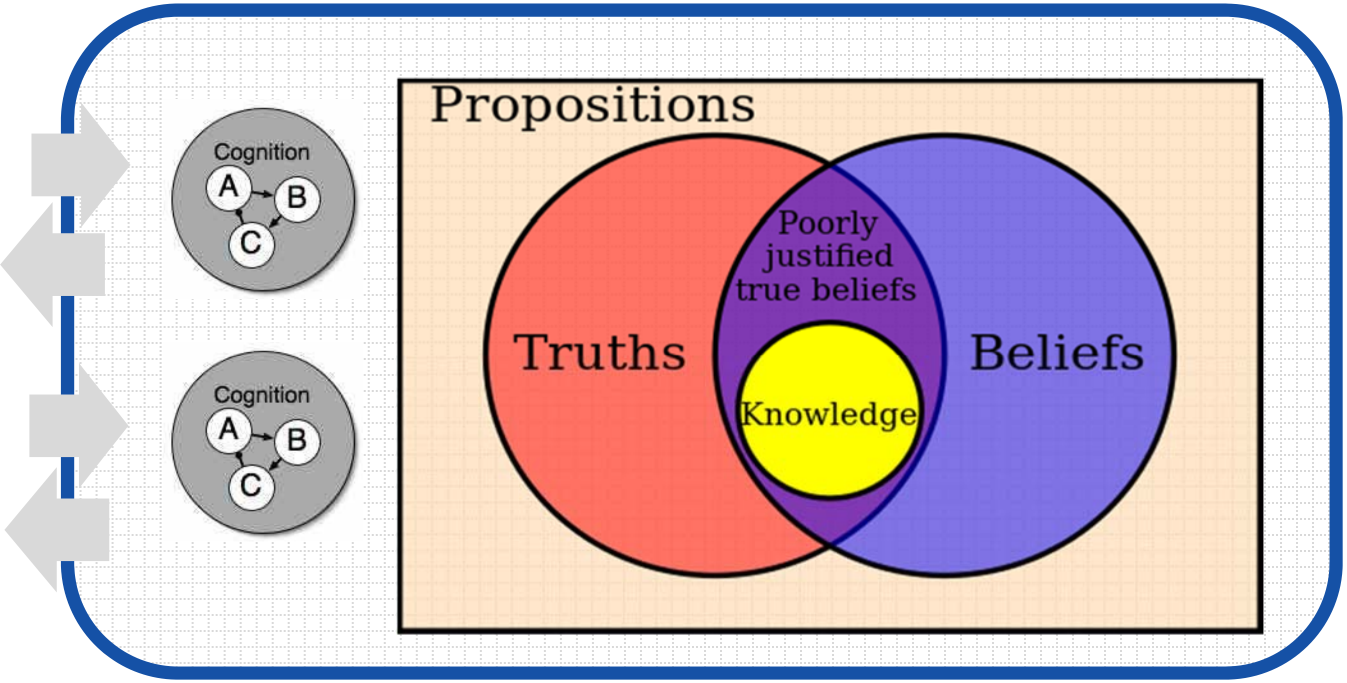

Concept of Operations
Verification and Validation
Operation and Maintenance
Project Definition
Requirements and Architecture
System Verification and Validation
Detailed Design
Integration, Test, and Verification
Project Test and Integration
Implementation
Time


Concept of Operations
Verification and Validation
Operation and Maintenance
Project Definition
Requirements and Architecture
System Verification and Validation
Detailed Design
Integration, Test, and Verification
Project Test and Integration
Implementation
Time

Knowledge-Base (KB) Strategy

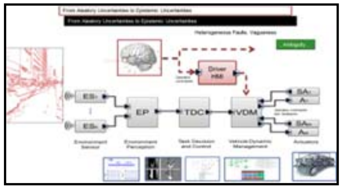
- A formal basis for describing, communicating and inferring
  - particular **operational truths** as well as
  - the **belief** and **knowledge** representing the awareness or comprehension of such truths

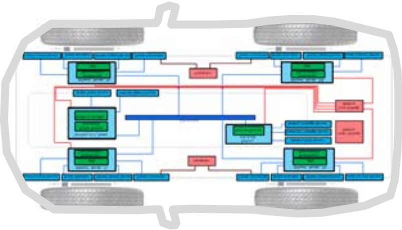


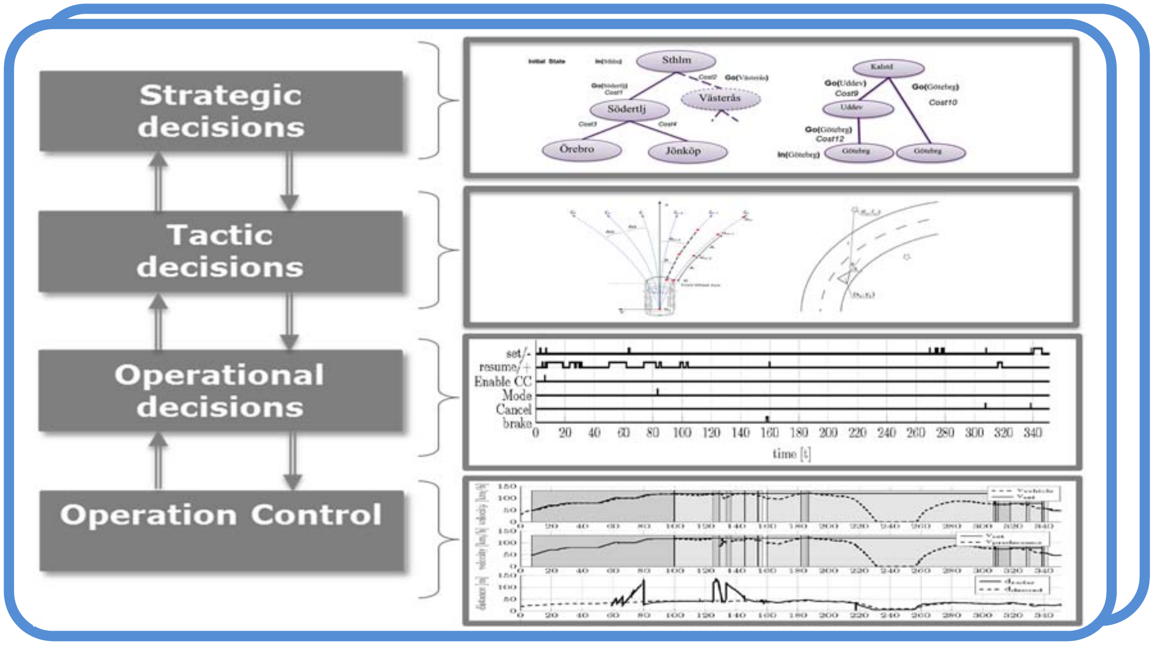


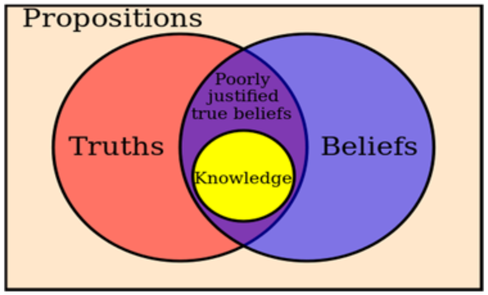
ADS Uncertainty

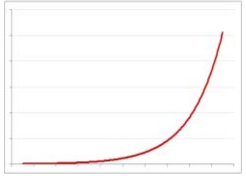
- **Aleatory uncertainty** - contextual complexity, e.g. unknowns due to emergent properties of traffic objects (i.e. Environment).
- **Epistemic uncertainty** - perception issue, caused by systematic unknowns caused by probabilistic algorithms, restricted observability, physical limitation, hidden variables, under-specification or semantic ignorant.
- **Capability uncertainty** - actual performance of a system
  - **Anomalies** i.e. the faults or errors exhibited by the computation and communication resources and vehicle plant, could result in additional nondeterminism of control functions.











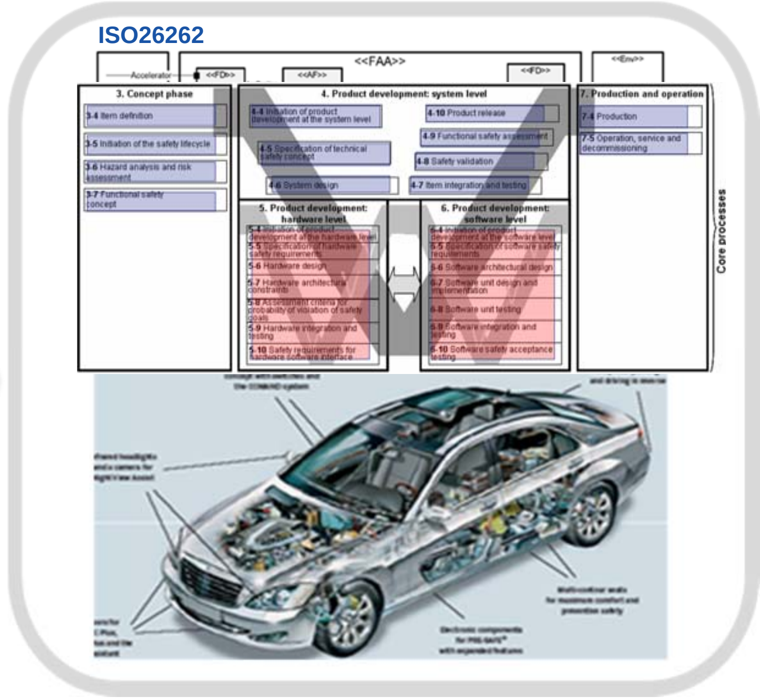
"Unjustified Belief"

Safety Agent

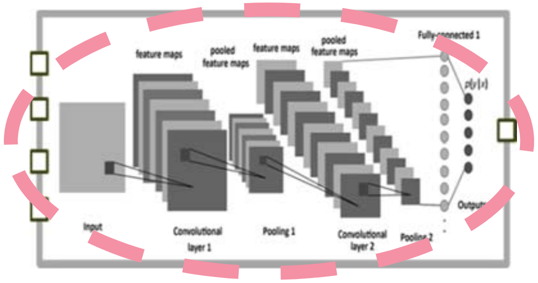
ADS Safety Engineering Challenges

Safety: Absence of unreasonable risks of causing damages to life, environment, or property

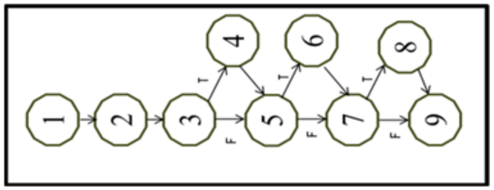
- Negative physical/chemical impacts (energy, harmful material)
- Level of risk acceptance (fault avoidance, removal and tolerance)


ISO26262
Accelerator
<<FD>>
<<AF>>
<<FAA>>
<<FD>>
<<Env>>
3. Concept phase
3-4 Item definition
3-5 Initiation of the safety lifecycle
3-6 Hazard analysis and risk assessment
3-7 Functional safety concept
4. Product development: system level
4-4 Initiation of product development at the system level
4-5 Specification of technical safety concept
4-6 System design
4-7 Item integration and testing
4-8 Safety validation
4-9 Functional safety assessment
4-10 Product release
5. Product development: hardware level
5-4 Initiation of product development at the hardware level
5-5 Specification of hardware safety requirements
5-6 Hardware design
5-7 Hardware architectural constraints
5-8 Assessment criteria for probability of violation of safety goals
5-9 Hardware integration and testing
5-10 Safety requirements for hardware software interface
6. Product development: software level
6-4 Initiation of product development at the software level
6-5 Specification of software safety requirements
6-6 Software architectural design
6-7 Software unit design and implementation
6-8 Software unit testing
6-9 Software integration and testing
6-10 Software safety acceptance testing
7. Production and operation
7-4 Production
7-5 Operation, service and decommissioning
Core processes

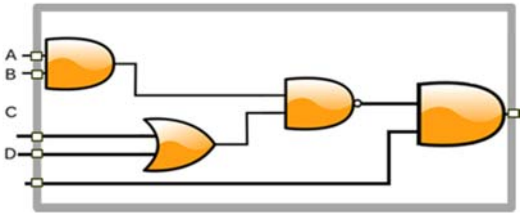
Trained behavior



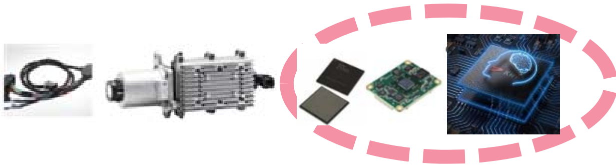
Coded behavior



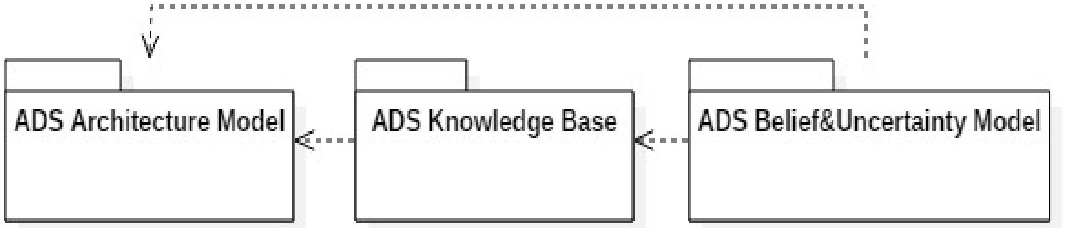
Wired behavior



fault avoidance, removal and tolerance



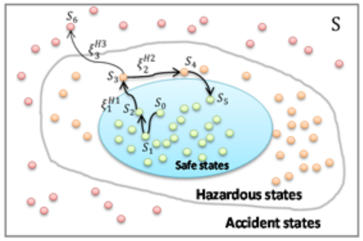
Integrating UM, KB, and EAST-ADL:

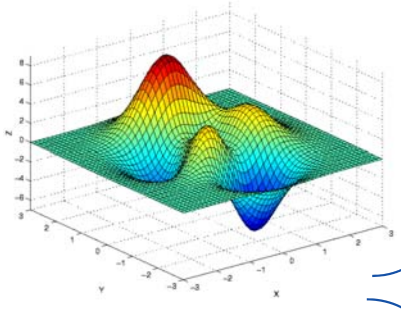


ADS Architecture Model

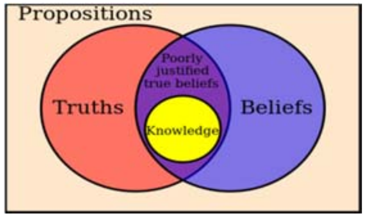
ADS Knowledge Base

ADS Belief&Uncertainty Model

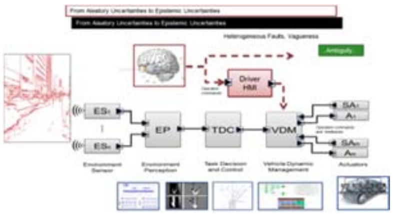


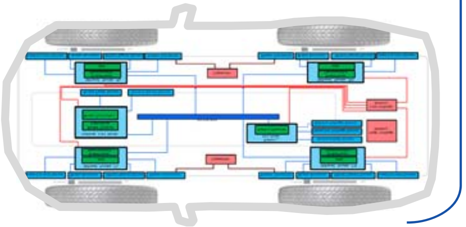


1. System (Operation) knowledge-Base

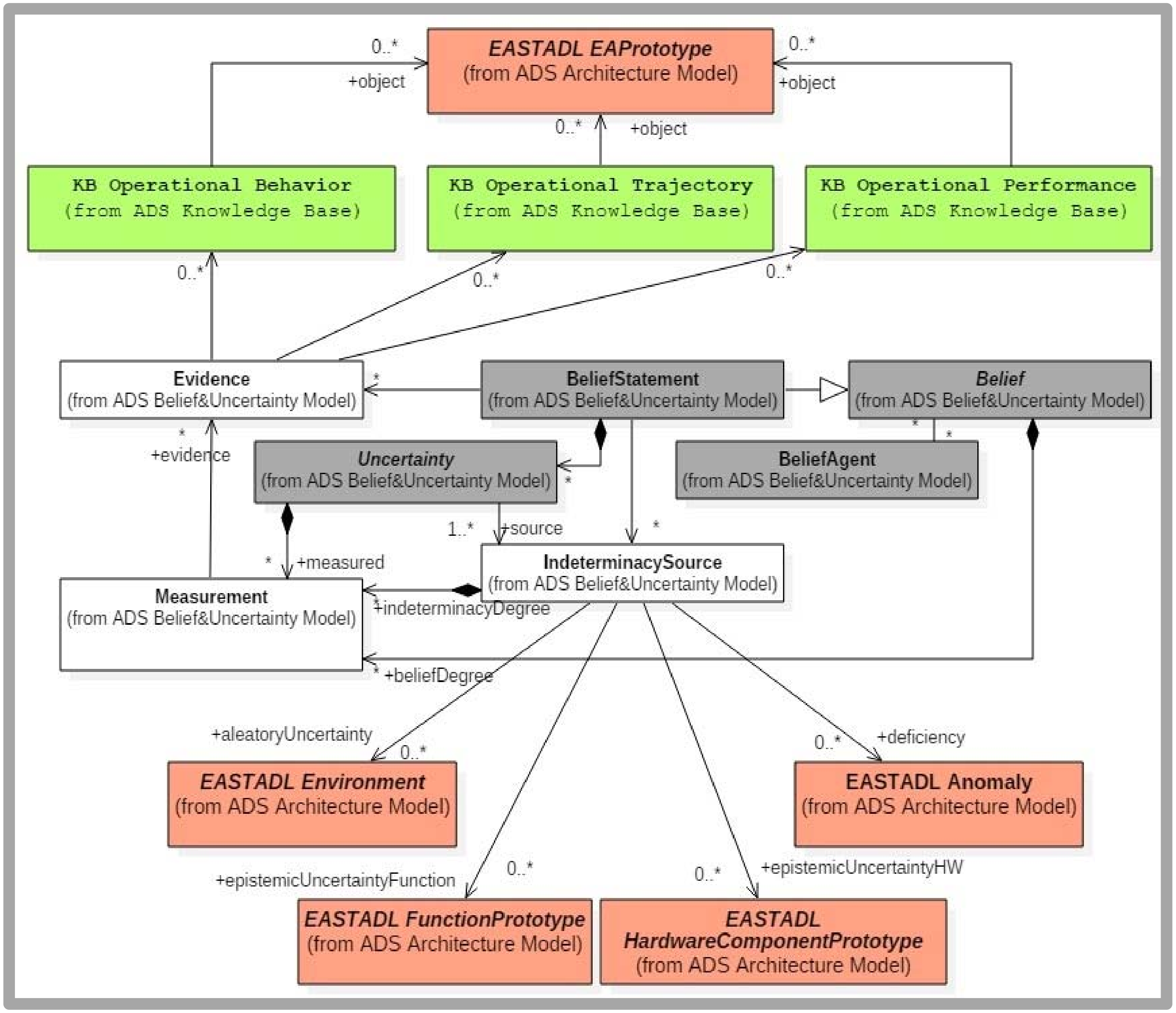


2. System Belief-Uncertainty Model





3. System Architecture



EASTADL EAPrototype
(from ADS Architecture Model)

KB Operational Behavior
(from ADS Knowledge Base)

KB Operational Trajectory
(from ADS Knowledge Base)

KB Operational Performance
(from ADS Knowledge Base)

Evidence
(from ADS Belief&Uncertainty Model)

BeliefStatement
(from ADS Belief&Uncertainty Model)

Belief
(from ADS Belief&Uncertainty Model)

Uncertainty
(from ADS Belief&Uncertainty Model)

BeliefAgent
(from ADS Belief&Uncertainty Model)

IndeterminacySource
(from ADS Belief&Uncertainty Model)

Measurement
(from ADS Belief&Uncertainty Model)

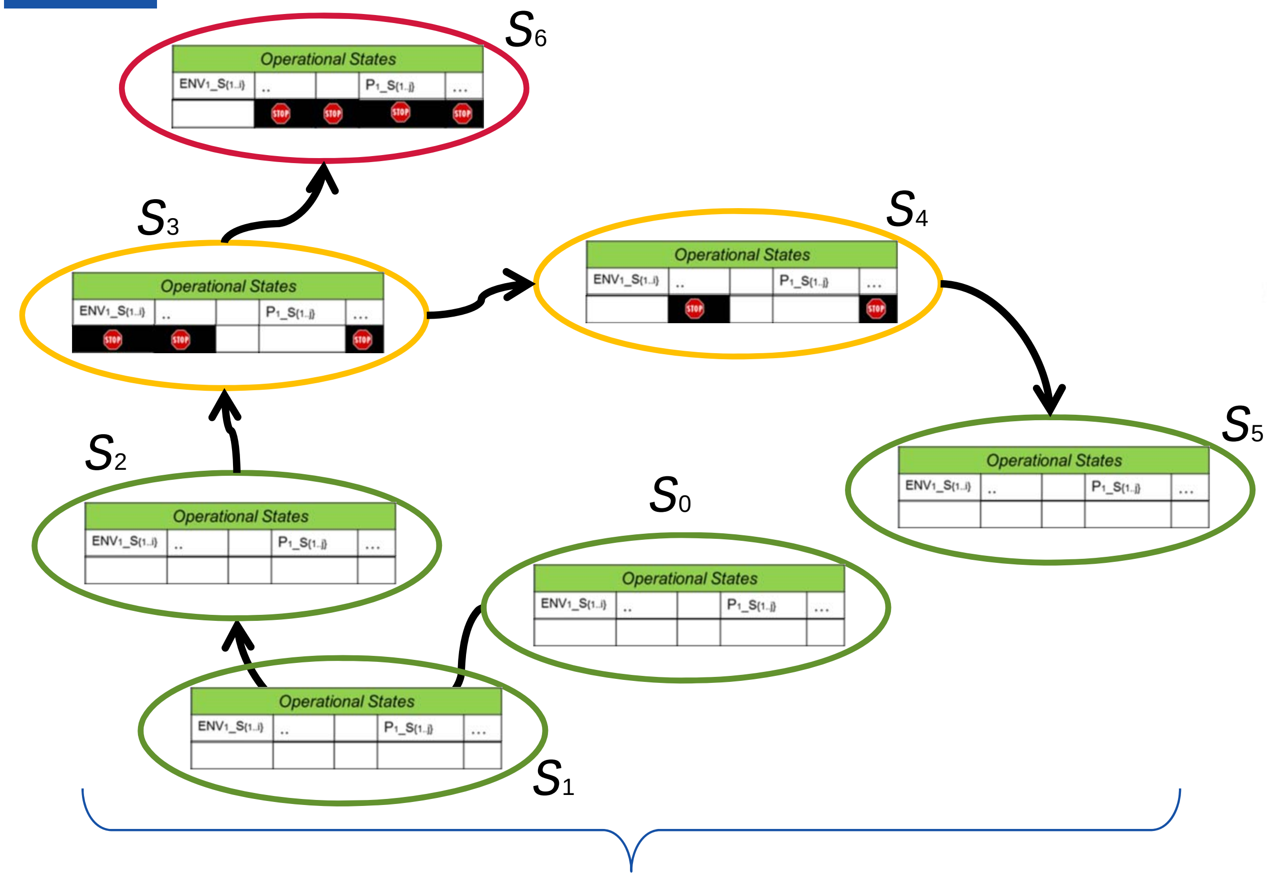
EASTADL Environment
(from ADS Architecture Model)

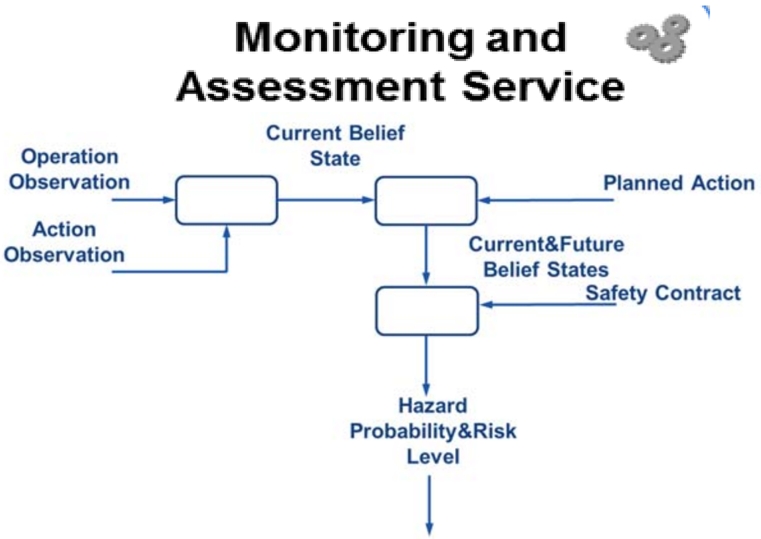
EASTADL Anomaly
(from ADS Architecture Model)

EASTADL FunctionPrototype
(from ADS Architecture Model)

EASTADL HardwareComponentPrototype
(from ADS Architecture Model)

Operation-Action Perception


S_6
Operational States
ENV1_S(1..i) .. P1_S(1..j) ...
S_3
Operational States
ENV1_S(1..i) .. P1_S(1..j) ...
S_4
Operational States
ENV1_S(1..i) .. P1_S(1..j) ...
S_5
Operational States
ENV1_S(1..i) .. P1_S(1..j) ...
S_2
Operational States
ENV1_S(1..i) .. P1_S(1..j) ...
S_0
Operational States
ENV1_S(1..i) .. P1_S(1..j) ...
Operational States
ENV1_S(1..i) .. P1_S(1..j) ...
S_1



Monitoring and Assessment Service

Operation Observation

Action Observation

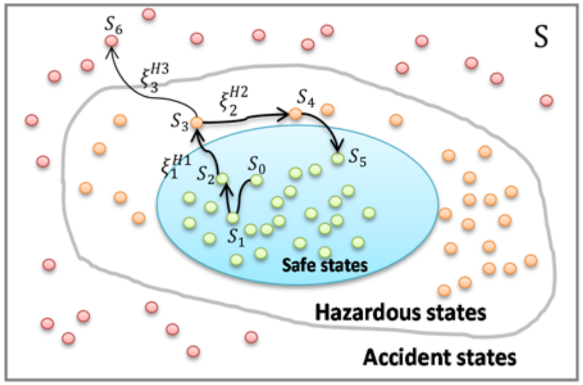
Current Belief State

Planned Action

Current&Future Belief States

Safety Contract

Hazard Probability&Risk Level



$$S_k = (S_{Env_k}, S_{Dri_k}, S_{Veh_k});$$

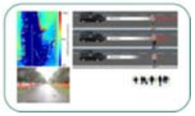
$$S_k \in S, S \subseteq S_{Env} \times S_{Dri} \times S_{Veh}$$

Uncertainty and Risk Inference

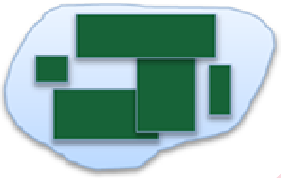
3 Perception performance (likelihood, TP)

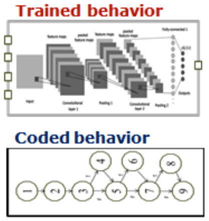
$$P(O_P|S) = \frac{P(S|O_P)P(O_P)}{P(S)} = \frac{P(S|O_P)P(O_P)}{P(S|O_P)P(O_P) + P(S|\neg O_P)P(\neg O_P)}$$

O_P = Object Pedestrian
S = Sensor Report TRUE

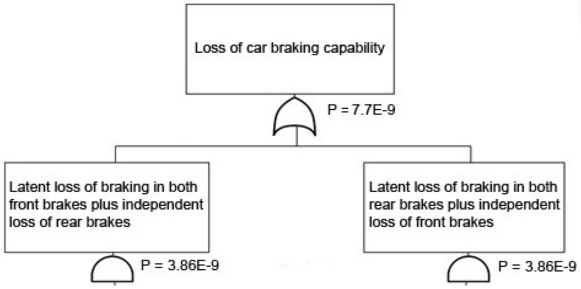


2 FORMAL VERIFICATIOI
(Real-world)

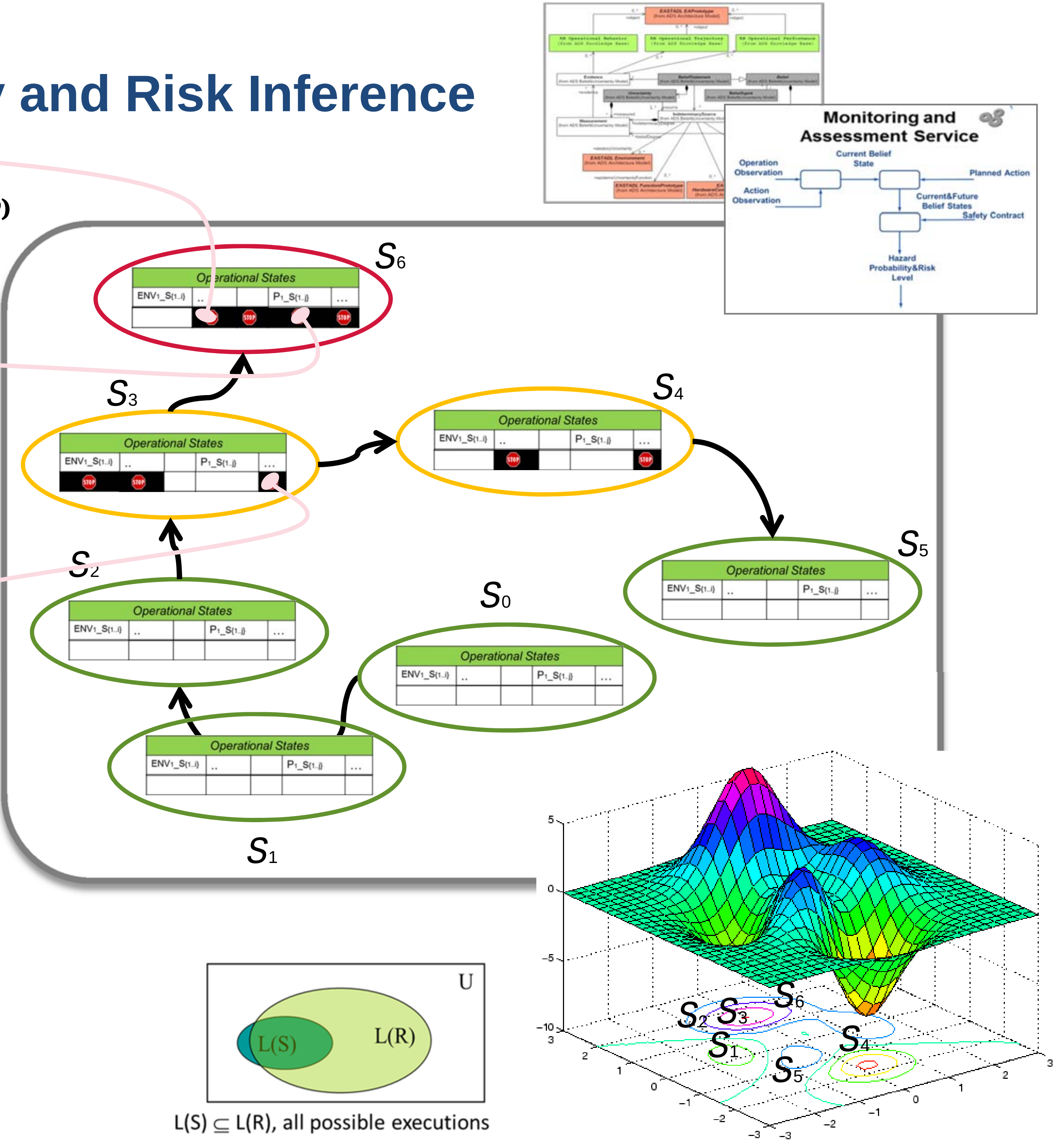




1 Brake failure (worst case by FTA)



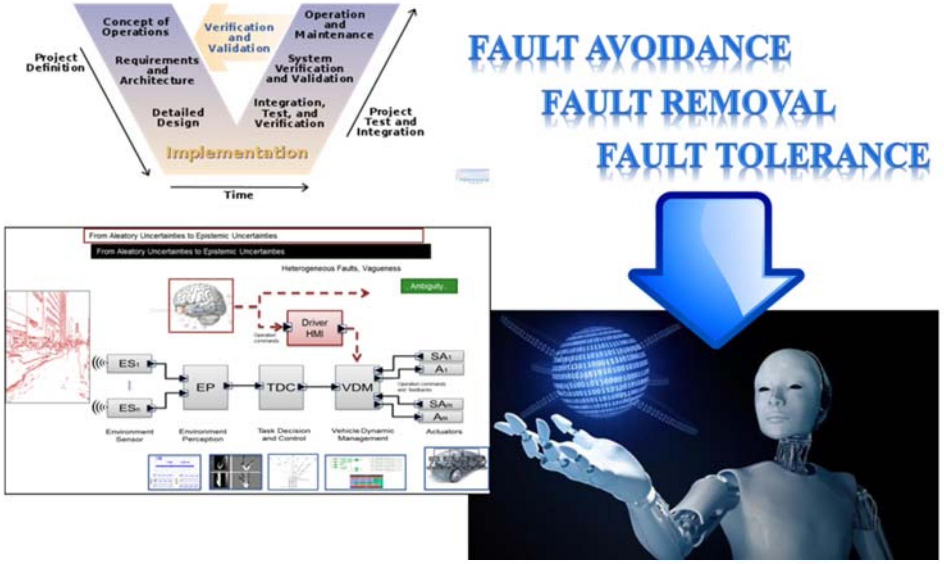
Loss of car braking capability
P = 7.7E-9
Latent loss of braking in both front brakes plus independent loss of rear brakes
P = 3.86E-9
Latent loss of braking in both rear brakes plus independent loss of front brakes
P = 3.86E-9



L(S) ⊆ L(R), all possible executions

Conclusion

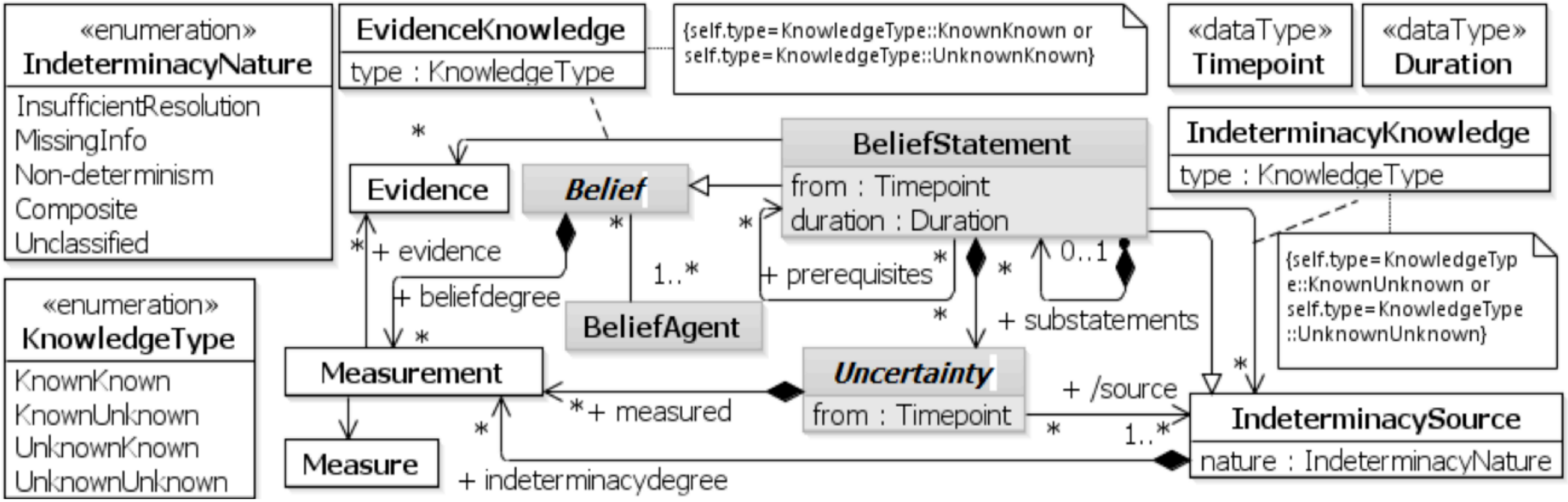
- AD exhibits uncertainties due to the operational contexts, the perception, computation and communication capacity.
- An knowledge-base constitutes the basis for an operation-action view
  - Requirement engineering, …
  - Safety engineering …
- For quality assurance, a paradigm shift in the engineering is needed for a systematic uncertainty management
  - Uncertainty modeling and probabilistic satisfaction assessment (residual risks)
  - Advanced safety “agent” for
    - state estimation and dynamic risk assessment
    - Knowledge enrichment and insurance cases


Concept of Operations
Verification and Validation
Operation and Maintenance
Project Definition
Requirements and Architecture
System Verification and Validation
Detailed Design
Integration, Test, and Verification
Project Test and Integration
Implementation
Time
FAULT AVOIDANCE
FAULT REMOVAL
FAULT TOLERANCE
From Aleatory Uncertainties to Epistemic Uncertainties
From Aleatory Uncertainties to Epistemic Uncertainties
Heterogeneous Faults, Vagueness
Ambiguity
Driver HMI
Operation commands
ES1
ESn
EP
TDC
VDM
SA1
A1
SAm
Am
Operation commands and feedbacks
Environment Sensor
Environment Perception
Task Decision and Control
Vehicle Dynamic Management
Actuators

UM for Belief State Modeling

Uncertainty Modeling

2. System Belief-Uncertainty Model

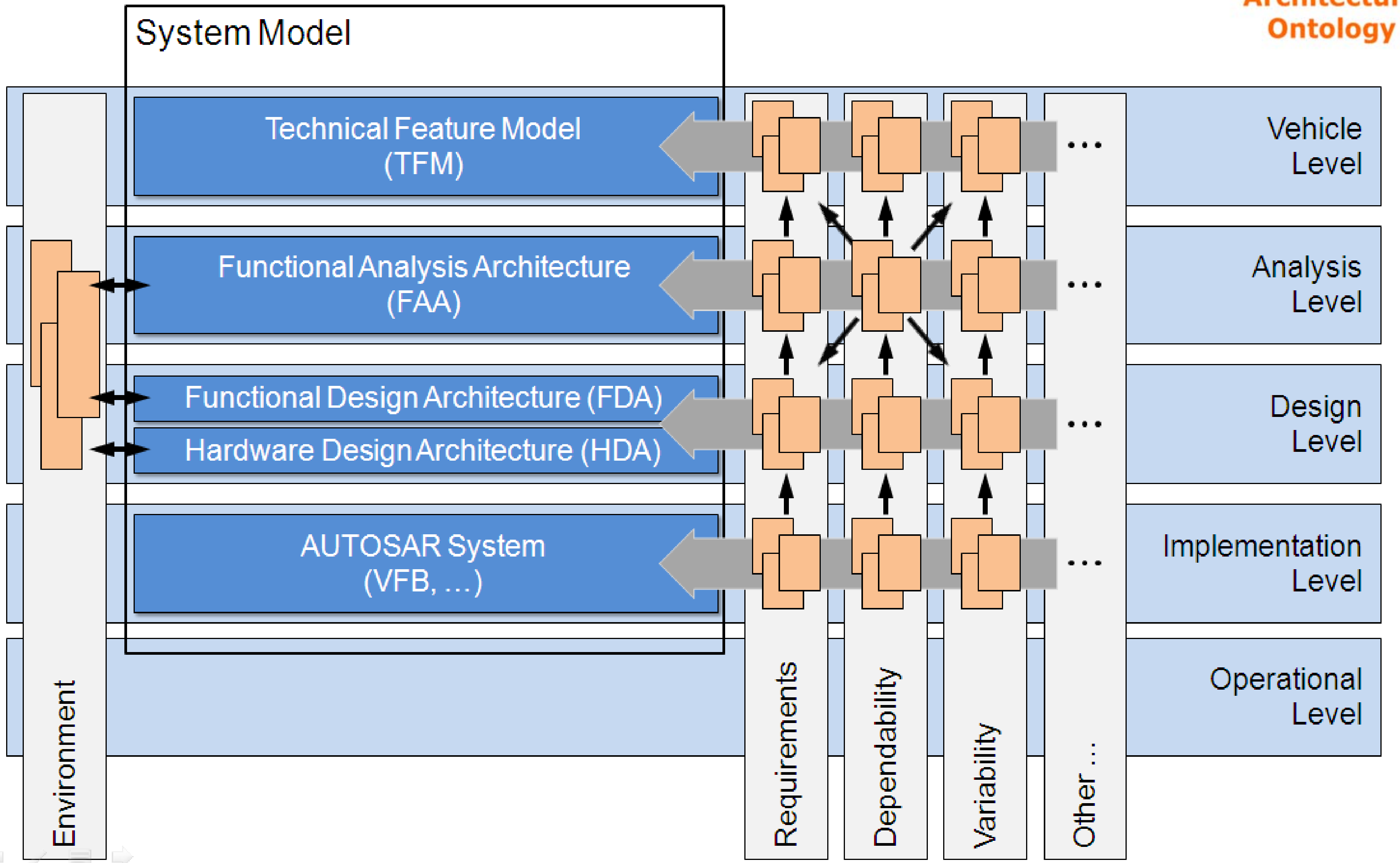


http://www.omgwiki.org/uncertainty/doku.php?id=start

- M. Zhang, B. Selic, S. Ali, T. Yue, O. Okariz, and R. Norgren, "Understanding Uncertainty in Cyber-Physical Systems: A Conceptual Model," presented at the ECMFA, 2016. Available: https://www.simula.no/publications/understanding-uncertainty-cyber-physical-systems-conceptual-model
- OMG. *Structured Metrics Metamodel* Available: http://www.omg.org/spec/SMM/

EAST-ADL – An Architecture Description Language (ADL)

3. System Architecture Ontology



http://www.east-adl.info/