Prof. Dr. rer. nat. Hermann Winner

Safe Driving Despite Uncertainties

Trotz Unsicherheiten sicher fahren
Steps in Safety Engineering

Avoidance of component failures of mainly mechanical parts

- Over-sizing
- Corrosion protection
- Life time design and maintenance free operation of a lot of components, which were replaced or maintained in the past
- Intensive component testing

Result: very small portion of accidents are caused by ‘technical failure’ in the total accident figure

- The weakest component: tires
- In general very high maturity, nevertheless some recall campaigns due to singular component problems occur
Steps in Safety Engineering

Passive Safety for reducing severity of accidents

- Restraint systems
- Stable passenger compartment with a stepwise energy absorbing body structure
- Marketing effect by consumer tests (e.g. EURO-NCAP)

Result: high reduction of fatality and injury rate per accident

- Potential seems be exhausted (at least for standard crash scenarios)
- Improvement only by support of active collision mitigation systems (Integral Safety) and/or individual adaptive restraint means
Steps in Safety Engineering

Active Safety for avoidance of accidents
(1. Step: Stabilisation support w/o. surroundings sensing)

- ABS ensures steerability and stability while braking
- ASR enables better traction and ensures steerability (frontal drive) and stability (rear axle drive)
- ESC (ESP) avoids additionally skidding in dynamic driving situations or critical pavements states.
- Brake assistant (mechanical, hydraulic, pneumatical) avoids unnecessary elongation of the braking distance

Result: Trend change of the total number of accidents

- Series equipment of new cars, high technology convergence
- Decisions on activation „mainly“ unambiguous, but driver interpretation (desired action threshold) difficult
Steps in Safety Engineering

Active Safety for avoidance of accidents
(2. Step: Assistance with surrounding Sensing)

- Informing systems (Park Distance Control, Lane Departure Warning, Rear Crash Warning, Lane Change Collision Warning) complement the human sensing and processing, but without „original“ damage potential
- Mild intervening systems (Adaptive Cruise Control, Lane Keeping Support) support actively the driver at his/her driving task within functional limits, and so reducing the potential for (re)actions not conform to the situation
- Systems with short, but strong interventions are able to reduce frequency or severity of accidents, but in the same way to produce critical states in cases, when the action is not conform to the situation.

Result: Market grows very dynamical after years of caution

- Potential to halve the number of injured persons in road traffic
Steps in Safety Engineering

Functional Safety

- Prevention of safety critical design errors in complex mechatronic functions
- Control of the development process, standards for assessment
- Current state: implementation phase of ISO26262 in the OEM and supplier companies

Result: ISO 26262 – a standard with some challenges for implementation

- ASIL-classification is still “somewhat” subjective
- Extension on others areas (e.g. PTW) needs relevant modifications
- Only failures or errors are treated, leading to nonconformity of the specification
- Uncertainty is not considered
Uncertainty

Characteristics:

- (Inherent) lack of knowledge about truth or correctness
- Leads to false negative or false positive state descriptions, or forged values of these
- Uncertainties are generated by incompleteness of knowledge. Instead of true states which are not accessible best guess model assumptions are used.
- These assumptions enable the function, but without achieving the same quality like true information sources.
- Leads to malfunction but without technical failures
Uncertainty

Comparison:

- Even the perception of a healthy human has still some uncertainties, e.g. optical illusions
Examples for uncertainty with surroundings sensing

Uncertainties depend on the processing step:

- **Feature detection**: quality depends on Signal-To-Noise-Ratio.
  - Despite electronic noise environmental conditions and signal processing artifacts affect the detection quality
- **Feature extraction** and **association** deliver the object information.
  - Neither false negative nor false positive errors are avoidable.
  - In addition: forged values for the objects may occur
- **Classification** decides about the relevance of the objects
  - True positive object, but false class

At the end sometimes incorrect results, although the system works conform to specification

- (Re-)Action on the basis „false“ data may not be conform to the situational need.
Description of Uncertainty by ROC

**Receiver Operating Characteristic**

- Variation of thresholds
- True-Positive-Rate versus False-Positive-Rate
- Benchmark for performance of classifiers
- Quality depends on the test data set
Description of Uncertainty according to Dempster-Shafer-Evidence-Theory

**Occupancy Grid**

- Segmentation of the surroundings on areal grid (cartesian or polar)
- In each cell two values between 0 (uncertain) and 1 (certain):
  - A: occupied; B: free (not occupied)
- Pseudo-probabilistic approach

Quelle oben: Grewe et al, Environment Modelling for Future ADAS Functions, 19th ITS World Congress 2012
Quelle unten: Grewe et al., ATZelektronik 7 #5, Umfeldmodelle: standardisierte Schnittstellen für Assistenzsysteme
Typical algorithm approaches for uncertainty

Probabilistic algorithms

- Bayesian networks
- Hidden-Markov networks

Diagram showing a Bayesian network with nodes X1, X2, X3, Y1, Y2, Y3, Y4, and edges labeled with probabilities.

Equations:

- \( P(W) = 0.5/0.5 \)
- \( P(B) = 0.6/0.4 \)
- \( P(G|W=ja) = 0.7/0.3 \)
- \( P(G|W=nein) = 0.1/0.9 \)
- \( P(T|B=ja,G=ja) = 0.9/0.1 \)
- \( P(T|B=ja,G=nein) = 0.5/0.5 \)
- \( P(T|B=nein,G=ja) = 0.5/0.5 \)
- \( P(T|B=nein,G=nein) = 0.1/0.9 \)

Sources:

- http://th.physik.uni-frankfurt.de/~mwagner/talks/Bayes.pdf, S.7

Legend:

- \( x \) — (verborgene) Zustände
- \( y \) — mögliche Beobachtungen (Emissionen)
- \( a \) — Übergangswahrscheinlichkeiten
- \( b \) — Emissionswahrscheinlichkeiten

Quelle: http://de.academic.ru/dic.nsf/dewiki/613738
Typical algorithm approaches for uncertainty

Probabilistic algorithms
- Bayesian networks
- Hidden-Markov networks

Pseudo-probabilistic algorithms
- Fuzzy Logic
- Neural Networks

Combinations

Typical applications
- Driver intention detection/interpretation
- Behavioral prediction/anticipation of other traffic participants

Result
- „Best supposition“, sometimes with uncertainty measure

Quelle:
Treatment of uncertainty in surroundings sensing (I)

Information processing

- Evaluation and optimisation on the basis recorded and annotated (labeled) test sequences
- Filtering and weighting considering multiple cycles (lead signal delay)
- Model based plausibility check (needs again model model assumptions)
- Minimizing mainly in the direction to low false positive actions
Treatment of uncertainty in surroundings sensing (II)

Functional design

- Design for Controllability:
  - Overriding by driver
  - Escape strategies in case of detected implausible states
  - Predictability of the actions by „smooth“ transitions, if compatible with functional goal

Result: Function below its potential, but safe

- Comfort, but also safety beneficial potential is not fully exhausted.
- The system safety is approved by its controllability.
Uncertainties and ISO 26262

Uncertainties may generate functional effects similar to „classical“ failure effects:

- Function is degraded
- Function has changed
- „Disturbances“ are generated

=> Adoption of the Hazard-Analysis for systems with uncertainty seems to be reasonable.

ISO 26262 addresses only systems with technical failures

- A methodological approach handling uncertainties in a functional safety process is missing.
Problems of ISO 26262 and ASIL-Logic in case of Driver Assistance Systems

ASIL-principle

- Risk classification is derived from Severity, Exposure and Controllability.
- From risk classification the probability of occurrence for non detected failure is determined, resulting in a very low frequency of potential critical cases.
- Actions not conform to the situation coming from uncertainties are in principle “not detected faults” and generate effects similar to failure effects.
- But: Frequent occurrence of such actions reduces the trust of drivers. So, the controllability will be increased due to learning of skills and the mental change.
- In case of actual technical failures with similar behavior like the above mentioned the controllability is improved and by this the failure effect will be less safety critical.
- Safety trade-off: Safety benefit vs. risk of the system
  - ISO 26262 treats only the risk, not the promised safety benefit
Validation of Advanced Driver Assistance Systems (I)

Functional tests

- In standards or regulations defined minimum requirements for function
  - Simple scenarios, simple criteria, e.g. no collision with test target
  - Ability to avoid false positive actions is not tested

Quelle: Bundesanstalt für Straßenwesen, eigenes Foto
Validation of Advanced Driver Assistance Systems (II)

Software-/Hardware-in-the-Loop-Tests

- Basis are recorded real raw data
- Huge effort for annotation
- Hardware changes and tolerances reduce the validity.
- For simulative stimulation appropriate models are missing. Neither for the environment nor for the signal propagation (e.g. Radar) suitable models exist.

Expectation

- Very effective and efficient validation method in the case of well suited test stimulation (but it is not the case for ADAS, currently)
- Doubts still remain whether the stimulation is sufficient.
Validation of Advanced Driver Assistance Systems (III)

Endurance and field tests

- Surroundings sensing is tested under realistic conditions (and data for offline evaluation can be acquired).
- Situation coverage is still incomplete even with 1 million km.
- Influence of drivers to the situation needs field tests with “normal” drivers. Even then a complete test coverage is impossible.

Results:

- “Classical” test results on hardware behavior and software errors
- Adaptation of thresholds for “optimal” compromise (best trade-off)
- Calibration of uncertainty models
Challenge Automated Driving

“Partially automated” driving

- More complex situation interpretation and more model assumptions
- Functional supervision by the driver
- Driver-in-the-Loop strategy
- ‘Design for Controllability’

Expectation

- Evolutionary development on the basis of existing ADAS => no relevant new methodological challenges for development and approval
- But more complexity and more effort
Challenge Automated Driving

“Highly automated” driving

- Supervision by the driver is no more required.
- Safe transition to driver when functional limit is reached.
- New challenge: Detection of functional limit in time for a safe transition (5 … 10 s in advance to transition).
  - Knowledge of the inherent uncertainty is required.
  - Driver’s ability for transition has to be ensured.
  - Machine situational and performance awareness

Expectation:

- Currently hard to imagine how the challenge can be addressed.
- Even in the case it would work: how would the approval procedure looks like?
Challenge Automated Driving

“Fully automated” driving

- Supervision by the driver is not required (like highly autom. driving)
- Transition to a risk minimum state when functional limit is reached.
- Strategy: Machine driving within a “risk corridor”
  - Knowledge of the inherent uncertainty form the driving style
  - Balancing of driving alternatives (always ready for a risk minimum driving maneuver)
  - Machine situational and performance awareness

Expectation:

- Currently hard to imagine how the challenge can be addressed.
- Even in the case it would work: how would the approval procedure looks like?
New challenges to validate and approve highly and fully automated driving

Self-dynamics

- Automated driving lives of its own in space and time. Just small differences leads to very different stories (like butterfly effect).
- Recordings loses its value for validation (loss might be reduced by segmentation and adaptive synchronisation)

Assessment measures

- Very often one will not find a single accident cause. The situation depends strongly on the previous history:
  - “Nearly every accident with collision partners would be avoided if one has started in traffic 10 minutes later. But this will not reduce the overall number of accidents.”
- For comparison to human driver a data base is not available, because only the “failed collision avoidances”, the accidents, were recorded.
Test Dilemma (I)

Macroscopic comparison with reference group

- Minimum requirement for systems with the objective to improve traffic safety:
  - Safer as reference group w/o this system
  - Accidents with damage to persons shall not be higher than within the reference group.
  - The number of accidents with only damage to property may be taken as indicator for extrapolation. But this is not valid in case of new systems.
### Test Dilemma (II)

**Today (Basis: ADAC statistics, 2010-2012):**

<table>
<thead>
<tr>
<th></th>
<th>Accident with damage to persons</th>
<th>Total mileage</th>
<th>Travel distance between to accidents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germany all road vehicles</td>
<td>300.000</td>
<td>$7.1 \cdot 10^{11}$ km</td>
<td>2.0 Mio. km</td>
</tr>
<tr>
<td>Germany passenger cars only</td>
<td>180.000</td>
<td>$6.0 \cdot 10^{11}$ km</td>
<td>3.3 Mio. km</td>
</tr>
<tr>
<td>Autobahn all vehicles</td>
<td>$18.000$ (Estimation from number of injuries)</td>
<td>$2.2 \cdot 10^{11}$ km</td>
<td>12.0 Mio. km</td>
</tr>
</tbody>
</table>

**in 10 or 20 years:**

- With increased field penetration of ADAS a doubling of the travel distance between two accidents with damage to persons can be predicted.
Test Dilemma (III)

Statistics result:

- In case of system twice as good as the reference a nearly 10x larger distance than the reference distance between two accidents is needed for 50% probability, to prove not to be worse than the reference on a level of significance of 5%.
- >100 million km
- Operation distance of 100 to 1000 passenger cars

Test travel distance/reference distance at 50%-success level
Challenge Uncertainty

Uncertainty is a third state between correct and false, significant more than usual measurement uncertainty coming from noise or bias.

- Undesired effects w/o occurrence of a failure.

- Treatment as trade-off between false negative and false positive mainly by experience values from the development people

- Accepted amount false positive actions strongly depend on controllability => Design for Controllability

Highly- and fully automated driving

- Uncertainty defines the function

- Requires uncertainty metrics

- Currently, nobody knows how to specify and to validate.
Conclusion

- ADAS and automated driving promise benefits for road traffic safety.
- Fundament of its functions is uncertain knowledge (particularly about vehicle’s surroundings and driver’s intention).
- *Design for controllability* warrants safety of the system, but limits the performance and potential safety benefit.
- Higher automation levels have to represent uncertainties as inherent part of their function.
- Methods for specification and validation have to be developed in accordance to the functional development. Otherwise the “approval trap” will stop the introduction.
- The ability to treat the uncertainties will be the key factor for progress of ADAS and driving automation.