Challenges in Achieving Explainability for Cooperative Transportation Systems: Perspectives From an Ongoing Research Project

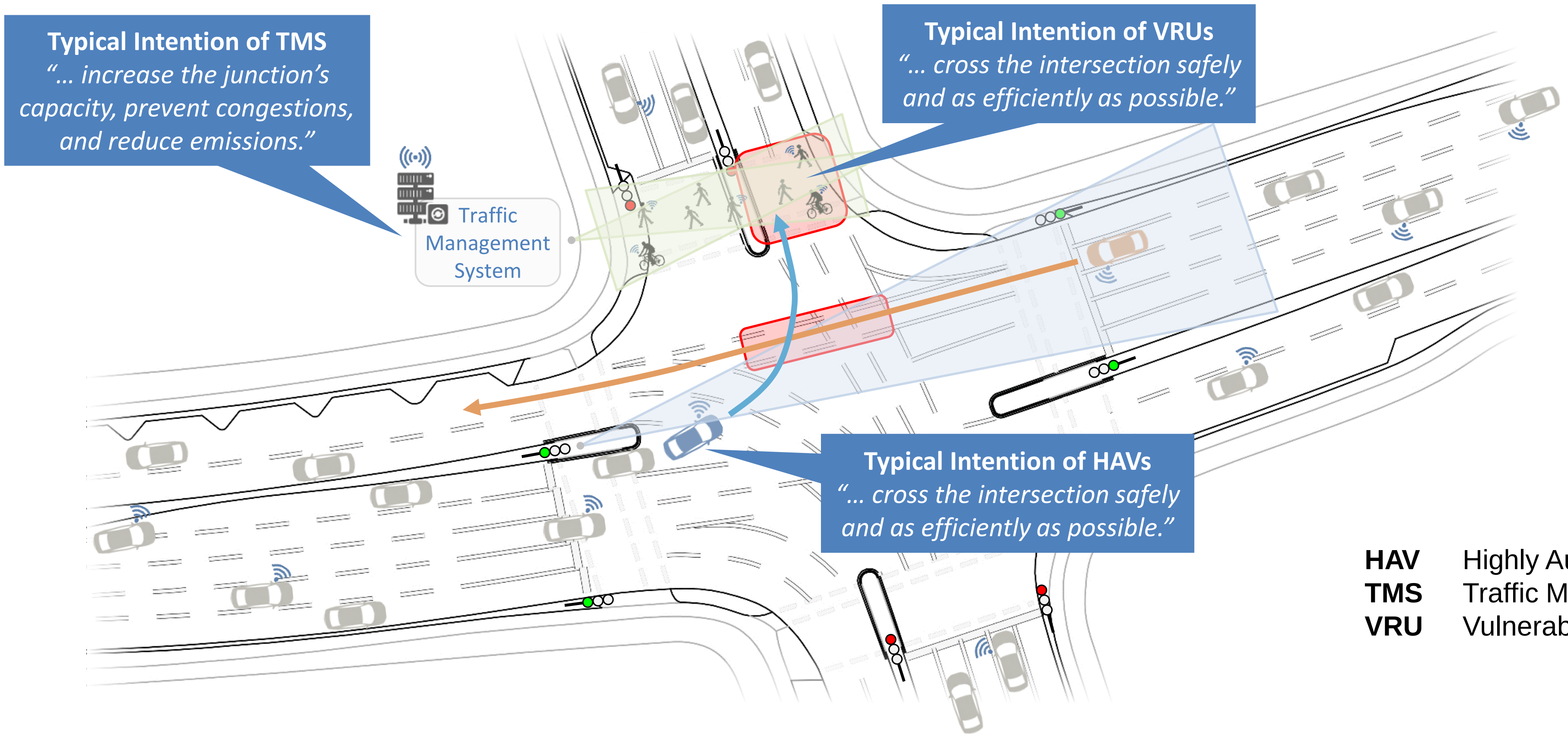
Björn Koopmann, Alexander Trende, Karina Rothemann, Linda Feeken,
Jakob Suchan, Daniela Johannmeyer, and Yvonne Brück

Björn Koopmann, M.Sc.
German Aerospace Center (DLR)
Institute of Systems Engineering for Future Mobility
Oldenburg, Germany

DLR

Cooperative Maneuvers at Urban Intersections

Use of Intelligent Infrastructure Systems to Support Highly Automated Driving



HAV Highly Automated Vehicle
TMS Traffic Management System
VRU Vulnerable Road User

[14] B. Koopmann, S. Puch, G. Ehmen, and M. Fränzle: *Cooperative Maneuvers of Highly Automated Vehicles at Urban Intersections: A Game-theoretic Approach*. In: *Proceedings of the 6th International Conference on Vehicle Technology and Intelligent Transport Systems (VEHITS'20)*. SciTePress, 2020, pp. 15–26.

Coordinated Cooperative Traffic with Distributed Intelligence (KoKoVI)
Ongoing, Large-scale Endeavor on Future Transportation Systems

- DLR-internal application-oriented research project
- Duration: 01/2022 to 12/2024 (1st phase) | Budget: 16.2 million €

- Main goal: **development**, **testing**, and **demonstration** of building blocks for **future cooperative transportation systems**
    - Idea: lift local coordination approaches to higher levels
    - Support for automated mobility services and remote operation
    - In-process testing in **traffic and driving simulations**
    - Final demonstration with **test vehicles** in real traffic

- Adoption of a **general view on explainability**
    - Results were derived directly from **practical experience**
    - Focus on different aspects of explanations and target groups
    - Identification of **challenges** in achieving explainability along the implementation of the KoKoVI use cases
    - Discussion of **selected approaches** to target these challenges

More information: https://verkehrsforschung.dlr.de/de/projekte/kokovi

Cooperative Transportation Systems

Lifting Local Coordination Approaches to Higher Levels

Users of Automated Mobility Services
- End users with mobile devices
- Logistics company staff

Traffic Participants
- Cyclists and pedestrians (VRUs)
- Human-operated vehicles (Non-HAVs)
- Highly automated vehicles (HAVs)
- Highly automated shuttle vehicles

Intelligent Transportation Infrastructure

Traffic Management Systems (TMS)

Disposition Systems

Remote Operation Center
- Remote operators
- Other technical staff


© Malte Helmhold


© DLR


© DLR


© Redd


© Marc Kleen


© Remy Gieling


© DLR


© DLR


© Janis Lucas

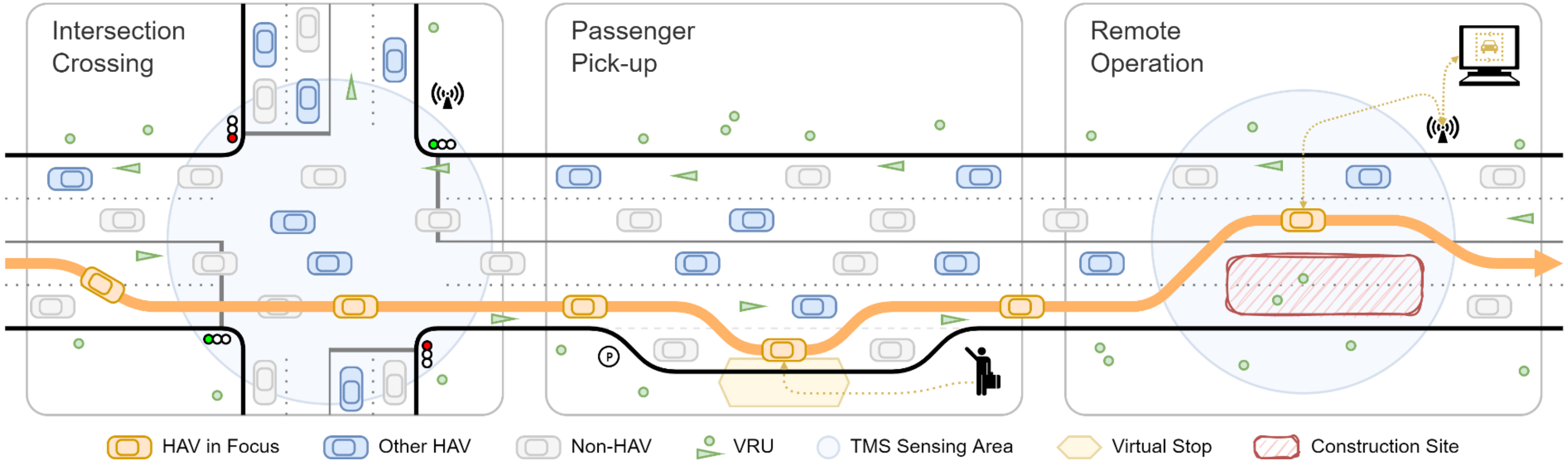

© Waldemar Brandt


© DLR


© DLR

Cooperative Transportation Systems

Example Scenarios Considered in KoKoVI (and Related Projects)



Intersection Crossing

- Support of an HAV by a traffic management system equipped with environmental sensors
  - Real-time information about the traffic situation in remote and poorly visible areas
  - Traffic light phases and future signal courses
  - Behavior and cooperation recommendations to increase traffic efficiency and safety
- Negotiation and execution of cooperative maneuvers such as lane changes and left-turns

Passenger Pick-up

- HAV in focus picks up a passenger at a "virtual stop" located in a parking bay
  - Processing of a trip request sent by the end user by a region-wide disposition system
  - Coordination of the shuttle fleet and selection of a shuttle vehicle to take over the ride
  - Initiation of the user's navigation to the stop
- Interaction with other HAVs, non-HAVs, and VRUs to enable a safe and efficient pick-up

Remote Operation

- Support of the HAV by a human remote operator in unknown or critical situations that cannot be handled by the vehicle's automation
- One of three possible countermeasures:
  1. Execution of minimal risk maneuvers
  2. Transfer of driving tasks to infrastructure
  3. Remote vehicle control by human operators
- Transmission of hazard warnings for oncoming traffic and following vehicles

Challenges in Achieving Explainability

Example Requirements From the Perspective of Four Key Roles

© DLR

Passengers of Highly Automated Vehicles

- Handing over control to an automated system requires a high degree of trust in technology
- Understanding the system is key for building trust
- Goal 1: provide passengers with the optimal amount of information about the system‘s behavior and intention
- Goal 2: prevent uncertainties due to deviating driving styles through systematic explanations

© Redd

Other Traffic Participants

- Explaining an HAV’s intention to other road users is important for the adoption of the technology
- Missing or incomplete communication can lead to misunderstandings, confusion, and mistakes
- Goal 1: indicate that HAVs are aware of the presence of non-HAVs and VRUs and anticipate their intended behavior
- Goal 2: communicate the HAV’s intention

© DLR

Remote Operators

- Monitoring HAVs requires constant vigilance
- Different user conditions such as under- or overload of cognitive demand can lead to vigilance decrements
- Goal: design HMIs such that remote operators are able to reliably detect and respond to critical systems states under various user conditions

© Malte Helmhold

System Developers and Integrators

- Establishing reliable interaction of systems with manufacturer-specific implementations of components and functions is a challenging task
- CTSs are built to provide value to society and should therefore be understandable and usable
- Goal 1: achieve explainability for human users
- Goal 2: provide error analysis and diagnosis, e.g., for legal proceedings after an accident

Challenges in Achieving Explainability

Identified Explainability Challenges and Selected Approaches

- Identification of 13 application-focused explainability challenges during system design
- Six generic challenges are relevant in all scenarios and from the perspective of all roles
- Additional challenges address specific properties of (single) systems and systems of systems

Generic Challenges

System Challenges

System of Systems Challenges

Group	ID	Title
I	**EC-1**	Different domain knowledge
	EC-2	Different levels of abstraction
	EC-3	Different visualizations
	EC-4	Different explanation goals
	EC-5	Different explanation paradigms
	EC-6	Different local goals and needs
II	**EC-7**	Missing explainability of black box components
	EC-8	Missing explainability of system context
III	**EC-9**	Explainability across (sub-)system boundaries
	EC-10	Manufacturer-specific implementations
	EC-11	Incompatible interfaces or communication protocols
	EC-12	Shared responsibility
	EC-13	Distributed decision making

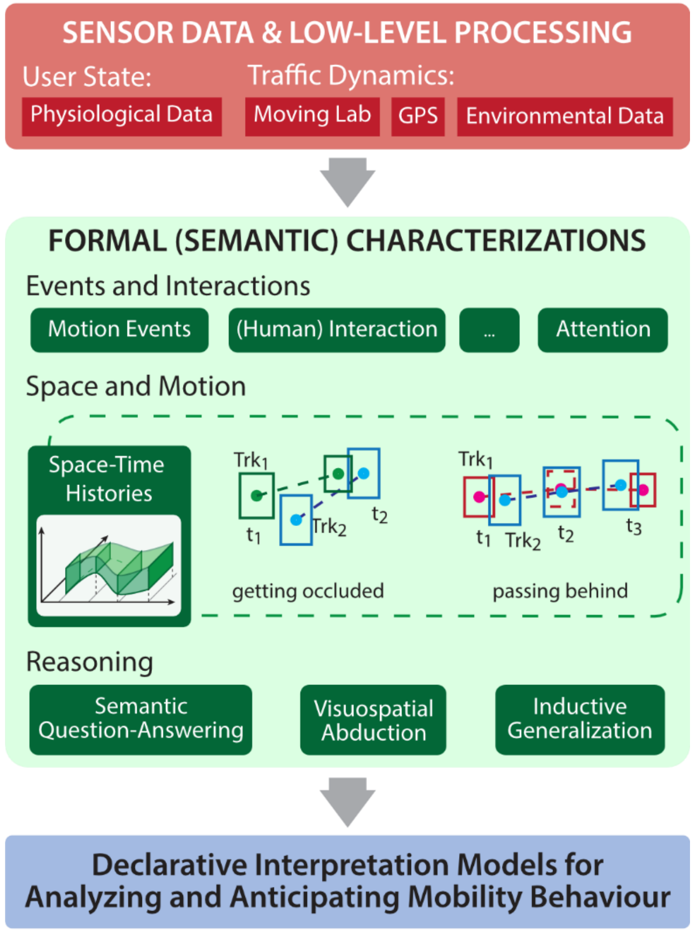
Solution Approaches

- Strategy Evaluation and Synthesis
- Explainability for Trustworthy Highly Automated Vehicles
- Explainable Interpretation Models
- Method for Designing Human-Machine Interfaces
- Architecture Development and Refinement
- Real-Time Handling and Safety Assessment

Developing Explainable Systems: Perspectives From KoKoVI

Explainable Interpretation Models

- Declarative characterizations for analyzing and anticipating mobility behavior
- Queryable relational models for the interpretation and projection of traffic dynamics and mobility behavior, focusing on:
    1. Encoding **high-level human-centered concepts** suitable for reasoning and learning about observed dynamics;
    2. **Grounding human interactions and behavior** with respect to low-level (subsymbolic) sensor data;
    3. Providing **human understandable abstractions** suitable to interface with and externalize inferred interpretations, i.e., for stating high-level **constraints** and **preferences**.
- Based on general and established (declarative) methods, e.g., Answer Set Programming (ASP)



[44] J. Suchan, M. Bhatt, and S. Varadarajan: *Commonsense Visual Sensemaking for Autonomous Driving: On Generalised Neurosymbolic Online Abduction Integrating Vision and Semantics*. In: *Artificial Intelligence*, Vol. 299, 2021.

Developing Explainable Systems: Perspectives From KoKoVI

Method for Designing Human-Machine Interfaces

Application and extension of the KONECT method (Ostendorp et al., 2016)

- Method for the development of safety-critical human-machine interaction interfaces
- Optimized to create visual components for fast and correct detection

- Problem: different user conditions such as under- or overload of cognitive demand can lead to vigilance decrements (Thomson et al., 2015)
  → Response times for critical event detection increase (Helton et al., 2011)

- Planned extension: consider impact of different user states on critical system state representation in the KONECT method to generate suitable HMIs and explanations for human remote operators

[46] M.-C. Ostendorp, T. Friedrichs, and A. Lüdtke: *Supporting Supervisory Control of Safety-critical Systems with Psychologically Well-founded Information Visualizations*. In: *Proceedings of the 9th Nordic Conference on Human-computer Interaction (NordiCHI'16)*. ACM, 2016, pp. 1–10.

Conclusion
Summary and Future Activities

Position Statement
- Achieving explainability in large-scale, cooperative transportation systems is a challenging task.
- The ability to explain processes and decisions is essential for such systems.
- From a systems engineering perspective, it is necessary to …
  - think explainability from the ground up,
  - handle explainability and trustworthiness as core properties of highly automated systems, and
  - carefully design these systems in such a way that the various needs for explanations are taken into account already in early design and engineering phases.

Contributions
- Derivation of 13 explainability challenges from practical experience
- Discussion of selected approaches being worked on in KoKoVI to target these challenges

Future Activities
- Re-evaluation of the identified challenges during system development
- Further reports on the progress in the development of solutions

Challenges in Achieving Explainability for Cooperative Transportation Systems: Perspectives From an Ongoing Research Project

Thank you for your attention!

Björn Koopmann, M.Sc.
German Aerospace Center (DLR)
Institute of Systems Engineering for Future Mobility
Oldenburg, Germany