Safe Control of Autonomous & Connected Vehicles (SCAV'17)*

Report from the 1st International Workshop at CPSWeek 2017

Mario Gleirscher
Technical University of Munich, Germany
mario.gleirscher@tum.de

Stefan Kugele
Technical University of Munich, Germany
stefan.kugele@tum.de

Jonathan Sprinkle
University of Arizona, Tucson, AZ
sprinkle@ece.arizona.edu

ABSTRACT

In this report, we summarize topics, challenges, and research questions discussed in the workshop contributions and during the sessions of our workshop. This summary has the purpose of leveraging the transfer of our findings into future activities of the *automatic vehicle control (AVC)* community.

Keywords

System safety, automatic vehicle control, autonomy.

1. ABOUT THE WORKSHOP

Our agenda included a keynote by Prof. Philip Koopman (CMU), five paper presentations, and an invited talk by Prof. Dan Work (UIUC). All contributions, submitted papers, and talk abstracts are published in the proceedings [1].

We conducted two break-out groups: one on the *assurance of adaptive control systems based on artificial intelligence (AI)*, and one on the *performance and safety of vehicular platooning*.

Finally, our panelists • *Prof. Philip Koopman (Carnegie Mellon University and Edge Case Research),* • *Akshaj Rajhans (MathWorks),* • *Prof. Raj Rajkumar (Carnegie Mellon University), and* • *Prof. Dan Work (University of Illinois at Urbana-Champaign)* shared their insights and opinions on the subject matter.

In the following sections, we give an overview of all subjects discussed in the workshop: core concepts, classes of systems (Sec. 2), their key objectives towards safe AVC (Sec. 3), and a list of research questions identified in this context (Sec. 4).

2. CONCEPTS AND SYSTEM CLASSES

Autonomy and Adaptive Control. Participants stayed agnostic of a specific definition of *autonomy*. We think it helps to view autonomy as a gradual concept: the more complex the control tasks to accomplish and the smaller the role of the human operator in the loop, the more autonomy we might attribute to an operational AVC system. Unsurprisingly, the safety of *adaptive optimal control* based on the use of AI techniques such as machine learning (ML) is of major concern under high degrees of autonomy.

System Classes. We can further conclude that *safe AVC* is desirable at multiple interrelated levels:

- At the level of **individual autonomous vehicles (AVs)**, our interest of **permissive safety** pertains to various degrees of automation, recently more often known as *highly (HAV) and fully automated vehicles (FAV)*.
- At the level of **connected AVs**, our interest lies on technologies connecting individual AVs with the infrastructure and, furthermore, with each other for various purposes.
- At the level of **cooperative AVs**, we focus permissive safety of platooning concepts.
- Finally, at the level of **hybrid traffic**, we regard permissive but safe mixing of manual and AV traffic.

We expect hybrid traffic to be one of the more relevant scenarios in the consumer car sector for the next decades. Traffic *solely consisting of AVs* was, hence, less of concern in our discussions.

3. SAFETY-RELATED OBJECTIVES OF AUTOMATIC VEHICLE CONTROL

In Tab. 1, we list several key objectives raised in the presentations, the break-out groups, and the panel. We mention these objectives grouped by the core qualities they mostly contribute to and the system parts or aspects they might have an impact on.

The objectives for core quality 3 represent a significant part of what is known as **safety of the intended function (SOTIF)**. Altogether, these objectives resemble the two perspectives of our workshop: (1) safety by designing safe system dynamics (cf. quality 3) and (2) safety by improving resilience, dependability, and security (cf. qualities 1 and 2c).

4. CHALLENGES AND OPEN QUESTIONS

Motivated by work of Koopman and Wagner [2], we provide an overview of the most relevant challenges and open questions raised in the workshop. Using the **id**entifiers in the first column of Tab. 1, we associate challenges and questions with the system parts, aspects, and objectives (Sec. 3).

Requirements. **Test data sets (5):** It is crucial to assess training data quality and to prepare representative samples. We have to use rich benchmarks (e.g. driving situation registers,[1] training data repositories) and experimentation platforms (e.g. [3]). Should industry be required to share test data, failure diagnosis data, controller source code, architectural knowledge?

Safety invariants (3a): We face the quest of minimum complexity at reasonable permissiveness. How well do simple but permissive obstacle detectors support reaching safe states? Is the denial of AVC features under unsafe conditions (e.g. bad weather)

*Accepted for publication in ACM SEN, Vol. 42:3, July 2017.

[1] E.g. `http://commonroad.gitlab.io`

Table 1: Key objectives in automatic vehicle control

Id.	Core Quality	of System Part or Aspect
1.	**Stability / Resilience**	of control algorithms
(a)	*Disturbance rejection / string stability*	of vehicle platoon controllers
(b)	*Fault-tolerance*	of control system architectures
(c)	*Attack-resilience / security*	of distributed controllers communicating in-vehicle, to-vehicle, and to-infrastructure
2.	**Performance / Efficiency**	at traffic level
(a)	*Permissiveness* (i.e. minimum constraints)	of safety invariants
(b)	*Continuity* (i.e. minimum congestion)	of traffic flow
(c)	*Availability*	of controllers and vehicles
3.	**Validity**	of controller requirements and designs
(a)	*Correctness and completeness*	of safety invariants, particularly, of requirements for AI-components
(b)	*Functional correctness*	of control software
(c)	*Correctness of quantization*	of signals
(d)	*Correct integration*	of humans in the loop
4.	**Accountability**	in AVC architectures
(a)	*Diagnosability*	of AI-based adaptive control
(b)	*Legibility*	of AI-components
5.	**Openness / Completeness**	of benchmarks, particularly, training data
6.	**Cost-effectiveness**	of assurance procedures

a feasible option? How can we specify permissive safety envelopes for AI-components? Where are the limits of reverse engineering of their objective functions?

Abstractions (3c): We have to improve the abstractions (of the control loop) implemented in safety monitors, particularly, assess residual uncertainty. What is the minimum required information to control safely in a specific operational situation?

Operational Features (3ab). Our discussions touched *unwritten*, culture-specific traffic rules. Moreover, how can human-driven vehicles and pedestrians be made aware of an AV's intentions? In addition, we have to safely balance permissiveness and defensive driving to increase traffic flow (2b). Furthermore, how can we build trust in vehicles when forming a **platoon** (1a)?

Human-in-the-loop adaptation (3d): We have to deal with hand-over from machine, take-over by machine, and even with scenarios of tele-operation.[2] Will tele-operation despite large feedback delays play any role for AV safety? Anyway, it will be important to keep human operators aware of an AV's current mode of operation.

Design, Architecture, and Optimization. Multi-stage monitor-actuators (e.g. for degradation) are practiced design options. Can we reduce their single-point faults with high impact (1b)? Do we have to achieve < 1 FIT[3] in adaptive or AI-based AVC? Can we reduce ML-uncertainty sufficiently (3c,4)?

[2]I.e. a supervisory control and data acquisition (SCADA) system for road traffic and vehicle operation.

[3]Less than one safety-critical failure in $10^9 hrs$ of operation.

Infrastructure data quality (1c,2b): We need to improve the sensing of the traffic state/flow. How much will AVs have to rely on the road infrastructure? Can we improve the corresponding data quality (e.g. traffic and geographical information)?

Security (1c): We have to identify unknown attack surfaces (e.g. intentional erroneous learning, indirect communication links) with high impact on AVC safety (e.g. jamming attacks to compromise platooning stability). Can we properly defend all such attacks?

Simplification (4): Will the simplification of currently used architectures drastically reduce costs and improve confidence in assurance? Can we avoid obstacles to other objectives in Tab. 1?

Assurance. **Increments, partial regression (3a,4):** Developers cannot be sure about what an AI-component has learned: How can we always find out whether, e.g., a deep neural network has learned unsafe behavior from a training data increment? How can we make learning monotonic w.r.t. safety? Hence, how can we improve legibility and diagnosability of AI-components?

Testing (5,6): We have to improve the testing of learned behavior. What is the minimum test coverage of AI-components to get sufficient confidence? What are the obstacles of viewing AI-components as black boxes and rather focus on the assurance of sophisticated safety monitors? How low is the current test coverage in the field? Which other methods will support certifiable assurance of HAF- and FAF-level AVC? Which problems can be solved by combining formally verifiable lookup-table approaches with off-line learning?

Politics and Education (1b,3,4,5). Who in society will define the *acceptable residual risk of AVs in hybrid traffic*? Does the "vaccination model" provide a meaningful hypothesis: safety improvements through AVs outweigh their negative side-effects? Will technical limits of AVs play a role in tightening their regulation? Will AV passengers need training in take- and hand-over scenarios? Will curricula in driving schools need to be adapted?

5. CLOSING REMARKS

Based on these first discussions, we are very much looking forward to continuing our community efforts and to tackling the most challenging issues as a scientific AVC community.

Acknowledgments. We take this opportunity to thank all the authors and presenters[4] for sharing their insights, our enthusiastic panelists for pointing at the peculiarities and "unknown unknowns", the members of the audience for posing many questions and actively taking part in the discussions, our program committee for careful reviews for the selection of contributions, and, last not least, the organizers and staffs of CPSWeek 2017 for their comprehensive support.

6. REFERENCES

[1] *SCAV'17: Proceedings of the 1st International Workshop on Safe Control of Connected and Autonomous Vehicles*, New York, NY, USA, 2017. ACM. `http://dl.acm.org/citation.cfm?id=3055378`.

[2] P. Koopman and M. Wagner. Challenges in autonomous vehicle testing and validation. In *SAE World Congress*, 2016.

[3] L. Paull, J. Tani, H. Ahn, J. Alonso-Mora, L. Carlone, M. Cap, Y. F. Chen, C. Choi, J. Dusek, Y. Fang, et al. Duckietown: an open, inexpensive and flexible platform for autonomy education and research. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2017.

[4]See the author list in the workshop proceedings [1].