Report from Dagstuhl Seminar 22102

Computational Models of Human-Automated Vehicle Interaction

Christian P. Janssen*[1], Martin Baumann*[2], Antti Oulasvirta*[3], Shamsi Tamara Iqbal*[4], and Luisa Heinrich†[5]

1　Utrecht University, NL. c.p.janssen@uu.nl
2　Universität Ulm, DE. martin.baumann@uni-ulm.de
3　Aalto University, FI. antti.oulasvirta@aalto.fi
4　Microsoft – Redmond, US. shamsi@microsoft.com
5　Universität Ulm, DE. luisa.heinrich@uni-ulm.de

──── **Abstract** ────

This report documents the program and the outcomes of Dagstuhl Seminar 22102 "Computational Models of Human-Automated Vehicle Interaction". At this Dagstuhl Seminar, we discussed how computational (cognitive) models can be used to model human-automated vehicle interaction. The seminar is motivated by developments in the field of semi-automated driving where humans and vehicles interact as teams to either both contribute to the drive (partnership) or to have safe transitions of control from vehicle to human and vice-versa. Computational (cognitive) models can be used in these situations to simulate or model human behavior and thought. Such models can be used among others to better understand human behavior, to test "what if" scenarios to guide design, or to even provide input to the vehicle about the human's potential behavior and thoughts. The seminar was attended by experts in various fields including computer science, cognitive science, engineering, automotive UI, human-computer interaction, and human factors. They represented academia, industry, and government organizations. With the attendees, we discussed five challenges of the field during panel discussion sessions:

- Challenge 1: How can models inform design and governmental policy?
- Challenge 2: What phenomena and driving scenarios need to be captured?
- Challenge 3: What technical capabilities do computational models possess?
- Challenge 4: How can models benefit from advances in AI while avoiding pitfalls?
- Challenge 5: What insights are needed for and from empirical research?

The attendees then split off into smaller working groups to discuss aspects of these challenges in more depth. Based on these discussions and other input from the attendees, this Dagstuhl report reports the following:

- an executive summary of the seminar
- position perspectives of all the attendees (section: "Talks")
- summaries of the various working groups (section: "Working Groups")
- summaries of the five panels (section: "Panel Discussions")
- an overview of relevant papers (section: "Open Problems")
- a research agenda with some of the most important developments and needs we identified for the field (section: "Open Problems")

All in all, we believe the seminar has shown that this field has lots of potential for development and an active community to tackle pressing issues. We can't wait to see what results the participants of the seminar will bring to the field in the future.

Seminar March 6–11, 2022 – http://www.dagstuhl.de/22102

────────────

* Editor / Organizer
† Editorial Assistant / Collector

1 Executive Summary

Christian P. Janssen (Utrecht University, NL)
Martin Baumann (Universität Ulm, DE)
Shamsi Tamara Iqbal (Microsoft – Redmond, US)
Antti Oulasvirta (Aalto University, FI)

This is the executive summary of Dagstuhl 22102: Computational Models of Human-Automated Vehicle Interaction, which took place March 6-11th 2022 in Hybrid format. The executive summary first summarizes the motivation of the seminar, then gives an overview of the broad challenges that were discussed, it then presents the results of the seminar. As this is only the summary, there are a lot more details about every item and result in other parts of this report, these are therefore referred to.

It has been a fruitful meeting, which sparked many research ideas. We want to thank all the attendees for their attendance and all the input they generate. We hope that it is of value to the community, and we can't wait to see what other results follow in the future based on discussions that started at this seminar!

Christian Janssen, Martin Baumann, Antti Oulasvirta, and Shamsi Iqbal (organizers)

Computational Models of Human-Automated Vehicle Interaction: Summary of the field

The capabilities of automated vehicles are rapidly increasing, and are changing human interaction considerably (e.g., [4, 6, 29]). Despite this technological progress, the path to fully self-driving vehicles without any human intervention is long, and for the foreseeable future human interaction is still needed with automated vehicles (e.g., [15, 22, 29, 37, 48, 47]). The principles of human-automation interaction also guide the future outlook of the European Commission [13, 14]. Human-automated vehicle interaction can take at least two forms. One form is a partnership, in which the human and the automated vehicle both contribute in parallel to the control of the vehicle. Another form is in transitions of control, where the automated system at times takes over full control of the vehicle, but transitions control back to the human when desired by the human, or when required due to system limitations. For both the partnership and the transition paradigm it is beneficial when the car and the human have a good model of each other's capabilities and limitations. Accurate models can make clear how tasks are distributed between the human and the machine. This helps

avoid misunderstandings, or mode confusion [45], and thereby reduces the likelihood of accidents and incidents. A key tool in this regard is the use of computational (cognitive) models: computational instantiations that simulate the human thought process and/or their interaction with an automated vehicle. Computational models build on a long tradition in cognitive science (e.g., [35, 36, 44]), human factors and human-computer interaction (e.g., [10, 39, 27], neuroscience (e.g., [12, 31]), and AI and engineering (e.g., [17, 42]). By now, there are a wide set of varieties that can be applied to different domains, ranging from constrained theoretical problems to capturing real-world interaction [38]. Computational models have many benefits. They enforce a working ethic of "understanding by building" and require precision in specification ([34], see also [8, 32, 41]). Models can test the impact of changes in parameters and assumptions, which allows for wider applicability and scalability (e.g., [2, 16, 44]). More generally, this allows for testing "what if" scenarios. For human-automated vehicle interaction in particular, it allows testing of future adaptive systems that are not yet on the road. Automated driving is a domain where computational models can be applied. Three approaches have only started to scratch the surface. First, the large majority of models focus on engineering aspects (e.g., computer vision, sensing the environment, flow of traffic) that do not consider the human extensively (e.g., [7, 18, 33]). Second, models that focus on the human mostly capture manual, non-automated driving (e.g., [44, 9, 25]). Third, models about human interaction in automated vehicles are either conceptual (e.g. [20, 22]) or qualitative, and do not benefit from the full set of advantages that computational models offer. In summary, there is a disconnect between the power and capabilities that computational models offer for the domain of automated driving, and today's state-of-the-art research. This is due to a set of broad challenges that the field is facing and that need to be tackled over the next 3-10 years, which we will discuss next.

Description of the seminar topics and structure of the seminar report

The seminar topics were clustered around five broad challenges, for which we provide a brief description and example issues that were discussed addressed. Although the challenges are presented separately, they are interconnected and were discussed in an integrated manner during the seminar. During the seminar, each challenge was discussed in a panel, with all attendees taking part in at least one panel. After each panel, the group was split up in smaller workgroups, and discussed the themes in more lengths. The summary of each panel discussion can be found later in this report under the section "panel discussions". The outcomes of the workgroups can be found later in this report under the section "workgroups". In addition, all attendees wrote short abstracts that summarized their individual position.

Challenge 1: How can models inform design and governmental policy?

Models are most useful if they are more than abstract, theoretical vehicles. They should not live in a vacuum, but be related to problems and issues in the real world. Therefore, we want to explicitly discuss how models can inform the design of (in-)vehicle technology, and how they can inform policy. As both of these topics can fill an entire Dagstuhl by themselves, our primary objective is to identify the most pressing issues and opportunities. For example, looking at:

- *Types of questions:* what types of questions exist at a design and policy level about human-automated vehicle interaction?
- *How to inform decisions:* How can models be used to inform design and policy decisions? What level of detail is needed here? What are examples of good practices?
- *Integration:* Integration can be considered in multiple ways. First, how can ideas from different disciplines be integrated (e.g., behavioral sciences, engineering, economics), even if they have at times opposing views (e.g., monetary gains versus accuracy and rigor)? Second, how can models become better integrated in the design and development process as tools to evaluate prototypes (instead of running empirical tests)? And third, how can models be integrated into the automation (e.g., as a user model) to broaden the automation functionality (e.g., prediction of possible driver actions, time needed to take over)?

Challenge 2: What phenomena and driving scenarios need to be captured?

The aim here is to both advance theory on human-automation interaction while also contributing to understanding realistic case studies for human-automation interaction that are faced for example by industry and governments. The following are example phenomena:

- *Transitions of control and dynamic attention:* When semi-automated vehicles transition control of the car back to the human, they require accurate estimates of a user's attention level and capability to take control (e.g., [22, 49]).
- *Mental models, machine models, mode confusion, and training and skill:* Models can be used to estimate human's understanding of the machine and vice-versa (e.g., [20]). Similarly, they might be used to estimate a human driver's skill level, and whether training is desired.
- *Shared control:* In all these scenarios, there is some form of shared control. Shared control requires a mutual understanding of human and automation. Computational models can be used to provide such understanding for the automation (e.g., [50]).

Challenge 3: What technical capabilities do computational models possess?

A second challenge has to do with the technical capabilities of the models. Although the nature of different modeling frameworks and different studies might differ [38], what do we consider the core functionality? For example, related to:

- *Compatibility:* To what degree do models need to be compatible with simulator software (e.g., to test a "virtual participant"), hardware (e.g., be able to drive a car on a test track), and other models of human thinking?
- *Adaptive nature:* Computational models aim to strike a balance between precise predictions for more static environments and being able to handle open-ended dynamic environments (like everyday traffic). How can precision be guaranteed in static and dynamic environments? How can models adapt to changing circumstances?
- *Speed of development and broader adoption:* The development of computational models requires expertise and time. How can development speed be improved? How can communities benefit from each other's expertise?

Challenge 4: How can models benefit from advances in AI while avoiding pitfalls?

At the moment there are many developments in AI that computational models can benefit from. Three examples are advances in (1) simulator-based inference (e.g., [26]) to reason

about possible future worlds (e.g., varieties of traffic environments), (2) reinforcement learning [46] and its application to robotics [30] and human driving [25], and (3) deep learning [17] and its potential to predict driver state or behavior from sensor data. At the same time, incorporation of AI techniques also comes with challenges that need to be addressed. For example:

- *Explainability:* Machine learning techniques are good at classifying data, but do not always provide insight into why classifications are made. This limits their explainability and is at odds with the objective of computational models to gain insight into human behavior. How can algorithms' explainability be improved?
- *Scalability and generalization*: How can models be made that are scalable to other domains and that are not overtrained on specific instances? How can they account for future scenarios where human behavior might be hard to predict [5]?
- *System training and corrective feedback:* if models are trained on a dataset, what is the right level of feedback to correct an incorrect action to the model? How can important new instances and examples be given more weight to update the model's understanding without biasing the impact?

Challenge 5: What insights are needed for and from empirical research?

Models are only as good to the degree as they can describe and predict phenomena in the real world. Therefore, empirical tests are an important consideration. Example considerations are:

- *Capturing behavioral change and long-term phenomena:* Many current computational models capture the results of a single experiment. However, behavior might change with more exposure to and experience with automated technology. How can such (long-term) behavior change be tested?
- *Capturing unknown future scenarios:* Many automated technologies that might benefit from computational models are not yet commercially available. How can these best be studied and connected to computational models?
- *Simulated driving versus real-world encounters:* To what degree are simulator tests representative of real-world scenarios (e.g., [19])?

Results

The seminar has generated the following results.

1. **Overview of state-of-the-art technologies, methods, and models.** The spectrum of computational modeling techniques is large [38, 21, 24]. Before and during the conference, we have discussed various methods and techniques. In particular, this report contains a dedicated chapter called "Relevant papers for modeling human-automated vehicle interaction" in which we report a long set of papers that the community identified as being relevant to the field. We encourage scholars to take a look at it.
2. **List of grand challenges with solution paths.** We have identified five grand challenges and discussed those in detail during the panels. Our chapters on "panel discussions" report the outcomes of these discussions. Moreover, the workgroups further report the in-depth discussions that smaller groups had about these challenges. See the section "working group" of this report. The results only start to scratch the surface of some of the grand challenges for the application of computational cognitive modeling that need to

be faced within the next 3 to 10 years, and their paths to solutions. Based on discussions, groups of authors plan to work on more papers and workshops around topics that they deemed worthy of further discussion. For example, we discussed whether there are specific driving scenarios that a computational model should be able to capture, and how success might be quantified (e.g., whether these challenges should take the form of competitions, akin to DARPA's Grand Challenge for automated vehicles [11] or "Newell's test" for cognitive models [3]).

3. **Research agenda to further the field.** This report also reports a research agenda that is intended to further the field. For each specific grand challenge, we have identified more specific areas of research that need futher exploration. We refer to the dedicated section in this report called "Research agenda to further the field". The organizers of the seminar will also organize a dedicated journal special issue around the topic, in which further results that arose from the seminar can be reported.

References

1   Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., Iqbal, S.T., and Teevan, J. (2019). Guidelines for human-AI interaction. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 1-13).

2   Anderson, J. R. (2007). How can the human mind occur in the physical universe? (Vol. 3). Oxford University Press.

3   Anderson, J. R., and Lebiere, C. (2003). The Newell test for a theory of cognition. Behavioral and brain Sciences, 26(5), 587-601.

4   Ayoub, J., Zhou, F., Bao, S., and Yang, X. J. (2019). From Manual Driving to Automated Driving: A Review of 10 Years of AutoUI. In Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (pp. 70-90).

5   Bainbridge, L. (1983). Ironies of automation. In Analysis, design and evaluation of man–machine systems (pp. 129-135). Pergamon.

6   Bengler, K., Dietmayer, K., Farber, B., Maurer, M., Stiller, C., and Winner, H. (2014). Three decades of driver assistance systems: Review and future perspectives. IEEE Intelligent Transportation Systems Magazine, 6(4), 6–22.

7   Brackstone, M., and McDonald, M. (1999). Car-following: a historical review. Transportation Research Part F: Traffic Psychology and Behaviour, 2(4), 181-196.

8   Brooks, R. A. (1991). Intelligence without representation. Artificial intelligence, 47(1-3), 139-159.

9   Brumby, D. P., Janssen, C. P., Kujala, T., and Salvucci, D. D. (2018). Computational models of user multitasking. Computational interaction design, 341-362.

10 Card, S. K., Moran, T., and Newell, A. (1983). The psychology of human-computer interaction. Hillsdale, NJ: L. Erlbaum Associates Inc.

11 Darpa (2020) The Grand Challenge. Accessed online on July 6 2020 at `https://www.darpa.mil/about-us/timeline/-grand-challenge-for-autonomous-vehicles`

12 Eliasmith, C. (2013). How to build a brain: A neural architecture for biological cognition. Oxford University Press.

13 European Commission (2018, 17 May). On the road to automated mobility: An EU strategy for mobility of the future (pp. 1–17). Brussels, BE. Communication COM(2018) 283 final.

14 European Commission (2020, 19 February). Shaping Europe's digital future. Brussels (BE). Communication COM(2020) 67 final.

15 Favarò, F. M. (2020). Unsettled Issues Concerning Semi-Automated Vehicles: Safety and Human Interactions on the Road to Full Autonomy. Technical report for the SAE. Warrendale, PA: SAE International. Retrieved from https://doi.org/s://www.sae.org/publications/technical-papers/content/epr2020001/

16 Gray, W. D. (Ed.). (2007). Integrated models of cognitive systems (Vol. 1). Oxford University Press.

17 Goodfellow, I., Bengio, Y., and Courville, A. (2016). Deep learning. Cambridge, MA: MIT press.

18 Helbing, D. (2001). Traffic and related self-driven many-particle systems. Reviews of modern physics, 73(4), 1067.

19 Hock, P., Kraus, J., Babel, F., Walch, M., Rukzio, E., and Baumann, M. (2018). How to design valid simulator studies for investigating user experience in automated driving – Review and hands-on considerations. Proceedings of the International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, 105–117. New York, NY: ACM Press

20 Janssen, C. P., Boyle, L. N., Kun, A. L., Ju, W., and Chuang, L. L. (2019). A Hidden Markov Framework to Capture Human–Machine Interaction in Automated Vehicles. International Journal of Human-Computer Interaction, 35(11), 947–955.

21 Janssen, C. P., Boyle, L. N., Ju, W., Riener, A., and Alvarez, I. (2020). Agents, environments, scenarios: A framework for examining models and simulations of human-vehicle interaction. Transportation research interdisciplinary perspectives, 8, 100214.

22 Janssen, C. P., Iqbal, S. T., Kun, A. L., and Donker, S. F. (2019). Interrupted by my car? Implications of interruption and interleaving research for automated vehicles. International Journal of Human-Computer Studies, 130, 221–233.

23 Janssen, C. P., and Kun, A. L. (2020). Automated driving: getting and keeping the human in the loop. Interactions, 27(2), 62-65.

24 Jeon, M., Zhang, Y., Jeong, H., P. Janssen, C.P., and Bao, S. (2021). Computational Modeling of Driving Behaviors: Challenges and Approaches. In 13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (pp. 160-163).

25 Jokinen, J.P.P., Kujala, T., and Oulasvirta, A. (2021) Multitasking in Driving as Optimal Adaptation under Uncertainty. Human Factors, , 63(8), 1324-1341.

26 Kangasrääsiö, A., Jokinen, J. P., Oulasvirta, A., Howes, A., and Kaski, S. (2019). Parameter inference for computational cognitive models with Approximate Bayesian Computation. Cognitive Science, 43(6), e12738.

27 Kieras, D. (2012). Model-based evaluation. In: Jacko and Sears (Eds.) The Human-Computer Interaction Handbook (3rd edition), 1294-310. Taylor and Francis

28 Kun, A. L. (2018). Human-Machine Interaction for Vehicles: Review and Outlook. Foundations and Trends in Human-Computer Interaction, 11(4), 201–293.

29 Kun, A. L., Boll, S., and Schmidt, A. (2016). Shifting Gears: User Interfaces in the Age of Autonomous Vehicles. IEEE Pervasive Computing, 32–38.

30 Levine, S. (2018). Reinforcement learning and control as probabilistic inference: Tutorial and review. arXiv preprint arXiv:1805.00909.

31 Marr, D. (1982). Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. San Francisco, CA: W.A. Freeman.

32 McClelland, J. L. (2009). The place of modeling in cognitive science. Topics in Cognitive Science, 1(1), 11-38.

33 Mogelmose, A., Trivedi, M. M., and Moeslund, T. B. (2012). Vision-based traffic sign detection and analysis for intelligent driver assistance systems: Perspectives and survey. IEEE Transactions on Intelligent Transportation Systems, 13(4), 1484-1497.

34 Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium.In Chase (ed.) Visual Information Processing. New York: Academic Press.

35 Newell, A. (1990). Unified theories of cognition. Cambridge, MA: Harvard University Press.

36   Newell, A., and Simon, H. A. (1972). Human Problem Solving. Englewood Cliffs, NJ: Prentice-Hall

37   Noy, I. Y., Shinar, D., and Horrey, W. J. (2018). Automated driving: Safety blind spots. Safety science, 102, 68-78.

38   Oulasvirta, A. (2019). It's time to rediscover HCI models. Interactions, 26(4), 52-56.

39   Oulasvirta, A., Bi, X., Kristensson, P-O., and Howes, A., (Eds.) (2018). Computational Interaction. Oxford University Press

40   Peebles, D., and Cooper, R. P. (2015). Thirty years after Marr's vision: levels of analysis in cognitive science. Topics in cognitive science, 7(2), 187-190.

41   Pfeifer, R., and Scheier, C. (2001). Understanding intelligence. Cambridge, MA: MIT press.

42   Russell, S., and Norvig, P. (2002). Artificial intelligence: a modern approach. Uppersaddle River, NJ: Pearson

43   SAE International. (2014). J3016: Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems. Warrendale, PA, USA: SAE International

44   Salvucci, D. D., and Taatgen, N. A. (2011). The multitasking mind. Oxford University Press.

45   Sarter, N. B., and Woods, D. D. (1995). How in the world did we ever get into that mode? Mode error and awareness in supervisory control. Human factors, 37(1), 5-19.

46   Sutton, R., and Barto, A. G. (2018). Reinforcement learning: An introduction. Cambridge, MA: MIT Press

47   Walch, M., Sieber, T., Hock, P., Baumann, M., and Weber, M. (2016). Towards cooperative driving: Involving the driver in an autonomous vehicle's decision making. In Proceedings of the International Conference on Automotive User Interfaces and Interactive Vehicular Applications, 261–268. New York, NY: ACM Press.

48   Walch, M., Mühl, K., Kraus, J., Stoll, T., Baumann, M., and Weber, M. (2017). From Car-Driver-Handovers to Cooperative Interfaces: Visions for Driver–Vehicle Interaction in Automated Driving. In G. Meixner and C. Müller (Eds.): Automotive User Interfaces: Creating Interactive Experiences in the Car (pp. 273–294). Springer International Publishing.

49   Wintersberger, P., Schartmüller, C., and Riener, A. (2019). Attentive User Interfaces to Improve Multitasking and Take-Over Performance in Automated Driving: The Auto-Net of Things. International Journal of Mobile Human Computer Interaction, 11(3), 40-58.

50   Yan, F., Eilers, M., Weber, L., and Baumann, M. (2019). Investigating Initial Driver Intention on Overtaking on Rural Roads. In 2019 IEEE Intelligent Transportation Systems Conference (ITSC) (pp. 4354-4359). IEEE.

2 Table of Contents

3 Overview of Talks

3.1 Computational modeling is key for advancing knowledge about human-AV interaction

Martin Baumann (Universität Ulm, DE)

The recent technological development is bringing the vision of automated driving within realistic reach. But for really exploiting the potential of this technology in terms of safety, efficiency and comfort human factors knowledge has to be considered and integrated into the design of these systems. Designing the interaction of humans both inside and outside with automated vehicles (AVs) is therefore a key factor for the technology's success. Unfortunately, we still lack important knowledge about the psychological mechanisms underlying the interaction of humans with AVs. At least in some cases, this is not due to the fact that we miss empirical data, but because we miss precise definitions and theories of the relevant mechanisms. This is where computational models of human-automated vehicle interaction come in and might play a key role in advancing the theoretical basis of our discipline in this field.

I see mainly three relevant phenomena that are essential for understanding human-AV interaction and that might profit from computational modelling:

i) how do humans construct and maintain an adequate comprehension of the current situation including system status,

ii) what are the long-term effects of interacting with AVs, and one aspect that is especially important here is how does trust into AVs evolve over time. Building computational models of trust and its development will definitely help us move forward regarding this supposed to be highly relevant but only loosely defined concept.

iii) how do humans interact and cooperate with each other in complex and realistic traffic scenarios. Current research results are mainly based on simple situations and one to one interaction scenarios. To tackle the complexity of such situations computational models allowing to simulate and to understand the processes in deeper level might be very helpful.

I hope that this Dagstuhl Seminar will bring us a step closer to solutions for these challenges.

3.2 Automated Driving and Cognitive Architecture Models

Jelmer Borst (University of Groningen, NL)

High-level cognitive architectures such as ACT-R can be used to simulate human driving under various circumstances [1]. Although they give a faithful characterization of the cognitive system, due to their rule-based nature they seem to be too brittle to operate an automated vehicle. Instead, I argue for four use cases:

1) understanding human cognition in driving;

2) a hybrid cognitive model-machine learning system, where the cognitive model informs the machine-learning part about expected behavior of other drivers and accompanying proper behavior;

3) a model-tracing approach, where the cognitive model is used to warn the driver when deviating from predicted behavior; and

4) a model-tracing approach where the cognitive model increases the automation level of the car when high workload is predicted for the driver.

References

1 Salvucci, D. D. (2006). Modeling driver behavior in a cognitive architecture. *Human Factors*, 48(2), 362–380. https://doi.org/10.1518/001872006777724417.

3.3 Reflections on 15 years of Modelling In-Car Multitasking Behaviour

Duncan Brumby (University College London, GB)

It was 15 years ago that I first wrote a computational cognitive model of driving. The model was used to make performance predictions for a range of different dual-task interleaving strategies that a driver could potentially adopt when entering a number to a mobile phone while driving. This work was published at CHI 2007 and it provided a great foundation for a lot of future research, including Chris Janssen's PhD thesis (who is now one of the organisers of this meeting).

At this meeting, I've been reflecting on three significant advances that have happened since 2007.

First, many modern cars have advanced driver support systems. This includes cameras that monitor the road so that the car itself can manage speed and lateral control. These systems have become quite wide spread, and show that the basic demands of the driving task are changing.

Second, mobile devices have got better. In 2007, the iPhone had only just been released. Now many cars have in-car displays to show route directions, and integrated voice user interfaces that can be used to play music or send messages. The secondary tasks that drivers can do are also changing.

Third, the basic science of modelling drivers has advanced. In 2007 we were just beginning to explore ways to use reinforcement learning methods to get our models to adapt strategies for driving and secondary task interactions. At the time we opted for a simple "black box" approach. Since this time the great AI Spring has delivered modern machine learning techniques that can be use by our models. The work being done by Jussi Jokinen and Antti Oulasvirta is now realising what we only hoped would have been possible 15 years ago. The models that we can develop are changing.

The aim of this meeting then is to reflect on the advances that we've seen in these three inter-related areas that are relevant for developing Computational Models of Human-Automated Vehicle Interaction: the driving task, the secondary task, and the basic science of cognitive modelling. I hope as a result of this meeting we can set out a research agenda for the next 15 years of change.

3.4 Modeling for AV-RU interaction

Debargha Dey (TU Eindhoven, NL)

Modeling an environment or human behavior, in essence, is looking at the world as a computer. But the world is complex, and within the dynamism we operate, it is often hard to quantify the specific dimensions of interest that provoke a certain behavior we observe. A complex world emerges from a complex interaction from an unquantifiable explosion of parameters, which tend to make any mathematical solution intractable. The job of a cognitive/ computation model is to simplify this quagmire of parameters and attempt to make some sense of this complexity with some informed guess, or in other words, illuminate the black box. Empirical research can help in systematically investigating the hierarchy of effect of the different parameters.

In the world of interactions within automotive human factors, we can say a lot about situations and negotiations, but we are not able to draw boundaries on problems and definitively say how to solve them. Furthermore, when we attempt to answer these questions through models, the expected reaction of the environment (e.g. driver, pedestrian, vehicle, etc.) are often not captured. Applying this to the context of my current research domain of eHMIs (external Human-Machine Interfaces) that facilitate AV-road user communication, linear models tell us how pedestrians will react and how a car should interact. However, the interaction from a car can cause further behavioral adaptations in a pedestrian, which current models are often unable to capture. The concept of interdependence – collaboration, coordination, and teamwork – is a grand challenge that seems to emerge as a gap in the state of the art.

Specifically in the context of eHMI research, the biggest potential benefit of a modeling approach emerges when tackling the problem of scalability. The state of the art has been mainly confined to a "one-car-one-pedestrian" setup in conducting empirical research. This is primarily because when investigating scalability of interactions, even at the minimum viable condition of testing scalability (i.e. just two pedestrians), the complexity introduced leads to the previously mentioned explosion of parameters, which make systematic empirical studies untenable. A potential approach is therefore to use a variety of theory- and naturalistic-data driven models to identify and decode the interactions in multi-agent scenarios, and conduct empirical studies with a more informed, limited set of parameters.

Another application for models is as a filtering tool. If a set of data does not fit a proven model derived from data or theory, it can be used as an indication of an outlier, which can be applied in practice as an alerting system. However, model interpretability is an important and non-trivial part of the equation if a model is to be used in that way: a machine-learned black-box model needs to be interrogated by a surrogate interpretable/ cognitive model. Furthermore, the question "how do we evaluate a model" stays wide open. What is the metric of a good model? And what is the appropriate thing to model for? The present? The future? A specific, imagined future? Points to ponder.

3.5 Data cautions in the use of machine learning tools

Birsen Donmez (University of Toronto, CA)

For some applications, like real-time driver state detection, the power of black-box Machine Learning (ML) tools is indisputable over other types of driver models. These ML tools enable the predictive modeling of complex nonlinear systems and can accommodate the modeling of a large number of predictor variables and their interactions. However, such complex models also require large datasets, which may or may not be directly available to researchers.

Caution should be taken when applying ML approaches to small datasets that empirical researchers traditionally collect. Although data may be collected at high frequency from each participant, experimental/observational scientific studies are generally limited in sample size (i.e., number of participants). Traditional statistical modeling techniques (model building and validation) were developed for small samples, whereas ML techniques assume large amounts of data. For example, researchers who have data from an experiment may have to split their training/test datasets within participants rather than across participants. But this approach can create data leakage resulting in ML models being rewarded for identifying participants rather than signals associated with the phenomenon of interest.

Caution should also be taken when researchers are lucky enough to obtain large datasets from other sources (e.g., OEMs). Real world data is messy, especially if there is not much control on how data is collected/selected for model training – leading up to examples such as facial recognition algorithms that are inequitable. Empirical researchers who design their own data collection are in a good position to apply their expertise and not fall into the same pitfalls that others do, by questioning how the data was collected, whether it is representative, and whether correlations exist in the data.

3.6 Triangulation of Cognitive Model and ML-based models

Patrick Ebel (Universität Köln, DE)

Out of the many extremely interesting topics that were discussed, what is of greatest interest to me is the question of how to combine cognitive models with learned (ML-based) models. How can we theoretical/mechanistic approaches improve machine learning-based approaches and vice versa? An effective combination of both could be the key to addressing the accuracy vs. Interpretability trade-off in HCI. In my opinion, an interesting way to go is to explore how complex interactions can be broken down into multiple subtasks/modules that only describe a small part of human behavior (glance allocation, pointing time. . .). These modules need to be restricted such that they only operate within certain mechanistic/cognitive boundaries. By doing so one can understand the different mechanisms of human behavior that in their combination describe a "complete" human-machine interaction without sacrificing accuracy due to abstract (but theoretically sound) models that lack prediction accuracy. However, even though this idea might be tempting, each module would still be trained in a specific context which leads to the problem of generalizability and raises the question of uncertainty prediction. I think that the effective usage of large natural data can be a solution to this problem. Leveraging large amounts of data we can represent different driving situations and contexts that would lead to an increase in generalizability.

3.7 Insights from Language Modelling Research

Justin Edwards (ADAPT Centre – Dublin, IE)

For decades, language models were rule-based and theory-driven. Language data is and has been plentifully accessible – models were not theory-driven in the 1960s because in-the-wild language data was not available, but because we didn't have the compute power to make use of large amounts of language data. The last decade has been very different for language models, with very large models like BERT and GPT-3 which, by utilizing tremendous power to train the models, modelling language in a quite complex, data-driven way. These data driven models have surpassed theory-driven models in many language tasks. They have been deeply flawed in other ways however, because of attempts to generalize data-driven models when the sample of language data is not general or ought not be generalized. This has resulted in applications, like chatbots that use racist language, or applications said to be capable of making moral judgments despite much of the training data coming from a Reddit advice forum. As new types of in-the-wild human behavioural data becomes available for contexts like human behaviour in SAE Levels 3 and 4 of driving, it is crucial that we examine the biases inherent to our data sources – the things that either will not or ought not generalize – and to remember that these biases can be caried through in our models. We must choose model architectures which are sensitive to biases in are data and we must be careful in how generally we try to apply models trained on idiosyncratic human behavioural data.

3.8 Towards a holistic architecture of human driver behavior

Mark Eilers (Humatects – Oldenburg, DE)

Intelligent design of human-automated vehicle interaction requires computational models of the human driving behavior and cognition. Particularly holistic models that can explain and predict multiple aspects of SAE Level 0 human driving behavior simultaneously seem equally important as modelling tool during the design of driving automation systems and their certification, and as a component of the system itself. Unfortunately, there is still a lack in commonly agreed upon models or architectures that can bear such a title.

As extensively discussed during the seminar, there is a common request for open-access and agreed-upon datasets, scenarios, benchmarks, and competitions, and first plans emerged to address this issue in the future. I'd argue that this endeavor should be accompanied by attempts to define (and provide software support for) an open-access architecture for a holistic model of human driving behavior that could evolve into a kind of gold-standard of driver models.

For a recent starting point, [1] provide a framework for a unified visuomotor model of a driver's lateral control during simulated driving that couples gaze and steering control in a three-layered architecture. It would be interesting to see this framework extended and harmonized with theories for e.g., longitudinal control, intention formation, route planning, and situational awareness, to name a few.

Such an architecture could start as a collection of (potentially competing) models that implement different aspects of human driving behavior (e.g., [2] two-point visual control model of steering) that could be plugged together and interfaced with the currently envisioned datasets, scenarios, benchmarks, and challenges in open-access driving simulators.

References

1 Lappi, O. and Mole, C. D. (2018). Visuomotor control, eye movements, and steering: A unified approach for incorporating feedback, feedforward, and internal models. *Psychological Bulletin*,144(10), pp. 981-1001. https://doi.org/10.1037/bul0000150

2 Salvucci, D. D. and Gray, R. (2004). A two-point visual control model of steering. *Perception*, 33 (10), pp. 1233-1248.

3.9 Reliable Autonomy Based on Imperfect Models

Martin Fränzle (Universität Oldenburg, DE)

From the perspective of design and safety analysis of highly critical human-cyber-physical systems, like automated driving at SAE levels 4 and 5, the demands on epistemic validity and accuracy of human models running as human state predictors in the car might seem prohibitive and the search for appropriate human models elusive. In my contribution to the panel, I made the point that this may not actually be true: if the models generate predictions which permit reliable (yet probably pessimistic) inferences in case of uncertainty then these models can seamlessly be embedded into safety-oriented system architectures when resorting to techniques of safe planning and control under uncertainty. Control theory as well as formal methods in computer science have in the past exposed various fundamental approaches to rigorous handling of uncertainties, which the human models and their corresponding execution mechanisms would then, however, need to be able to interface to.

These approaches rely on comprehensive mathematical representation of uncertainties concerning system state and evolution, building on possible world semantics in that they represent a system by sets or distributions of possible states at any point of current and extrapolated future time. This may seem a minor variation from the currently prevalent state-based models, given that the latter frequently come in stochastic variants which implicitly define the required distributions. The interfacing problem, however, results from the execution mechanisms employed for computing state extrapolations over these models: simulation-oriented mechanisms generating state traces, even if they do so using means of randomized simulation for stochastic models, are not suitable, as approximating a distribution via massively iterated randomized simulation is grossly inadequate for embedding in hard real-time into, as necessary for an online mechanism. What we instead need are mechanisms directly computing distributions in a time-resolved manner.

We argue that using such mechanisms, quite unreliable or, more precisely, uncertain human models -in experiments we have used such with just 61% predictive accuracy- could effectuate significant safety gains in embedded control applications. Technically, the human model here acts inside the embedded control as an online proxy of the human. This proxy enables predictive evaluation of the consequences of possible control actions, thus permitting the technical system to pursue rolling-horizon model-predictive control that is provident to the human.

Such an online embedding of human models would, however, require that the models to be embedded feature rigorous real-time guarantees, that they can incorporate the (uncertain) state evidence provided by in-situ measurements, that they can exploit the latter for computing best-possible estimates of current state, and that the provide time-resolved distributional state predictions over reasonable horizons of the imminent future. All these are prerequisites for interfacing seamlessly to rolling-horizon model-predictive control. Current models or rather their execution mechanisms fail to meet these by primarily being targeted to simulation, i.e., to trace generation rather than to state estimation by distributions.

3.10 How to model situation awareness and predict driver take-over ability

Luisa Heinrich (Universität Ulm, DE)

With higher levels of automation, the human's role shifts from actively steering the vehicle to passively monitoring the system. However, as long as fully automated driving (SAE L5) is not available, humans will have to intervene in the driving task at times. Scenarios in which transitions are likely to occur should therefore receive special attention. For example, when exiting a highway, or when entering urban traffic. Starting with SAE L3 of automation, the human driver may disengage from the driving task and perform other non-driving related activities. When a TOR is issued, he or she may be out-of-the-loop and incapable of understanding the situation and acting appropriately to traffic events, especially in safety-critical situations with small time-budgets. In order to support the human during transitions in automated driving, we need to predict the driver state. Does the human have enough situation awareness to safely take over control of the vehicle? With situation awareness, I don't mean being able to consciously reproduce knowledge of the current situation, but having a well enough understanding to make appropriate driving decisions – implicit rather than explicit knowledge. Computational models may inform us about driver state by considering metrics relevant to the build-up of situation awareness, for example eye-tracking (behavioral) or physiological data (e.g., heart rate). The question arises as to what data could or should be fed to the model in order to make valid predictions about the state of the driver and his or her ability to take over the driving task, and how the model can be validated if the variable of interest, the implicit understanding of the situation, is not measurable.

3.11 Hybrid models – integrating Machine learning approaches into cognitive models

Moritz Held (Universität Oldenburg, DE)

A key step to understanding driving behavior are the underlying cognitive processes while driving. With the help of computational cognitive models, we aim to not only better understand points of human failure but also provide theory-driven predictions for different

driving scenarios. Combining theory-driven white box models like ACT-R models with data-driven machine learning techniques ("black box models") is an intriguing approach towards creating predictive models which are usable in autonomous vehicles. In the Dagstuhl Seminar we identified key areas in which black box approaches are more useful either from a computational point of view (e.g., lower computation time in highly time-critical scenarios) or when it comes to estimating Driver states/traits. On the other end, white box approaches can be useful when the model needs to be robust to handle unseen scenarios or when the aim of the model is to understand the cognitive processes responsible for the behavior. Hybrid models, which integrate machine learning techniques to restrict the behavior of a cognitive model seem promising.

3.12 The Nomadic Worker: Autonomous Vehicles as Future Worksites

Shamsi Tamara Iqbal (Microsoft – Redmond, US)

With a rapid transition to a new future of work, the traditional definitions of where, how and when people can get their best work done is continuously evolving. An important finding of the years of the pandemic showed that people appreciated not having to commute as they had to work remotely and found that they could rather use the time previously spent on commute more productively while working from home. Yet with hybrid work practices, many workers will be returning to some form of commute at least during some of their work week. A key research question is how workers can effectively use the commute time – be it for their productivity or personal wellbeing needs without jeopardizing driving safety. In hybrid work scenarios, this becomes an even more important question, as efficient use of time for personal and collaborative productivity and wellbeing will be paramount for workplace performance. I propose a research agenda looking at the challenges of making the car a temporary worksite, where people can safely and comfortably get work done or attend to their personal needs without the worry of "wasting time in commute". I envision three stages – 1) fully manual (present day), 2) Level 3 where the driver is still in charge of driving with some autonomy support from the car, 3) fully autonomous – and the nature of work that people can get done will be very different from one stage to the next. The value of having three stages is that we can start experimenting right away and inform vehicles of the future of the general challenges of working in the car – limited attention span, environmental constraints such as motion sickness, lack of large workspaces – and extend and adapt to the unique needs of a particular stage. This work cross cuts both the domain of Future of Work, as well as Autonomous Vehicles and will facilitate innovations of how to get work done in non-traditional worksites.

3.13   An exciting time for the field of Cognitive Modeling of Human-Automated Vehicle Interaction

Christian P. Janssen (Utrecht University, NL)

It is an exciting time for the field of computational cognitive models of automated driving, and that excitement is palpable at this Dagstuhl meeting — even when attending it remotely. The field of cognitive modeling has been around for at least five decades (depending on how you count, but for example going by this commentary of Allen Newell [1]), and has seen a lot of progress. The last few years have seen a growth in the types of AI and modeling techniques that have been developed for and in Human-Computer Interaction [2]. Moreover, there is excitement to include these techniques in the Automotive domain [3]. A challenge is though that with the breadth of techniques, and the breadth of application areas it is easy to get lost in translation. Following a preceding Dagstuhl [4], we had already identified that it is valuable to distinguish between simulations of the agent, its environment, and scenarios [5]. Similarly, at this Dagstuhl we had various discussions on how to push the field forward: refine and integrate techniques for modeling into the automotive domain. My hope and expectation is that it will be a win-win-win situation. First, through modeling the automotive field will gain better insights into its users and can better design for it. Second, the social sciences and related fields will have an area to test their models, to see if they also work in more practical sessions and to identify where refinement is needed. Third, through this endeavor a solid scientific community will form. I can't wait to see what the future brings!

References

1 Newell, A. (1973) You can't play 20 questions with nature and win: projective comments on the papers of this symposium. In Chase, W.G. (Ed.) *Visual Information Processing*. New York: Academic Press.

2 Oulasvirta, A. (2019). It's time to rediscover HCI models. *Interactions*, 26(4), 52-56.

3 Jeon, M., Zhang, Y., Jeong, H., P. Janssen, C., & Bao, S. (2021). Computational Modeling of Driving Behaviors: Challenges and Approaches. In *Extended Abstracts of the 13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 160-163).

4 Riener, A., Boll, S., & Kun, A. L. (2016). Automotive user interfaces in the age of automation (Dagstuhl Seminar 16262). In *Dagstuhl reports* (Vol. 6, No. 6). Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.

5 Janssen, C. P., Boyle, L. N., Ju, W., Riener, A., & Alvarez, I. (2020). Agents, environments, scenarios: A framework for examining models and simulations of human-vehicle interaction. *Transportation research interdisciplinary perspectives*, 8, 100214.

3.14   Technical capabilities for computational models

Myounghoon Jeon (Virginia Polytechnic Institute – Blacksburg, US)

Computational models are used to simulate and predict human behaviors by quantifying the relationship between research parameters and behavioral outcomes. In this process, precision of prediction and efficiency of the modeling process might be necessity for modeling, rather than capabilities. Once these fundamental components are satisfied, we can consider adaptability or expandability of modeling, which we can call compatibility. The first technical capability the model should have for compatibility would be sensitivity. It is about whether the model can be sensitive to changes in resource demand. The next one is selectivity. It is about whether the model can be sensitive only to differences in specific resource demand. For example, if task difficulty can be resolved by multiple resources (e.g., secondary task while driving), but the model has only visual components, it might not be able to describe the entire human behavior. Then, the research question would be whether the model can be expanded by having other resources, such as auditory (speech and non-speech) components. The next capability for compatibility would be diagnosticity. It is about whether the model can indicate when human behaviors vary and can indicate the cause of variation. For example, the cognitive architecture framework may not capture other types of human states–e.g., emotions, fatigue, trust, mind wandering while driving. So, the research question would be whether the model can be combined with other constructs to predict and explain other than cognitive constructs. Finally, in terms of precision of the model, as long as the model can postulate the same stimulus-response mapping, it should work. This is the original meaning of ecological validity. Therefore, the simulation does not necessarily have high fidelity software and hardware compatibility (which is more about external validity), but the important question would be whether the model can have the psychophysical similarity. These technical capabilities (sensitivity, selectivity, diagnosticity, and ecological validity) will be useful to assess the compatibility of computational models in future.

3.15   HMI design for autonomous vehicles: methodologies and intercultural analyses

Xiaobei Jiang (Beijing Institute of Technology, CN)

To ensure driving safety, efficiency, and comfort in mixed traffic environment with autonomous vehicles (AVs), these vehicles must be able to understand the plans and states of partner road users, and can communicate and coordinate their actions with them. With the increasing number of AVs being tested and operating on roads, external HMI (eHMIs) are proposed to facilitate interactions between AVs and other road users. Taking eHMI for crossing pedestrians as an example, many methods like user interview with images & videos, Wizard of Oz (WoZ), virtual reality, the Delphi method, on-road experiment, etc have been conducted in studies focusing on the effect of eHMIs. But different methods are subject to different biases, even yielding conflicting results. This problem should be carefully considered. As there are cultural differences in the way how goals and states are expressed by HMI as well as how HMI is interpreted by others, it is important to understand these differences to allow the integration of AVs into different cultural contexts.

3.16   Benefits and Challenges of Computational Cognitive Models

Jussi Jokinen (University of Jyväskylä, FI)

The benefit of using computational cognitive models is that they force the modeller to make theoretical and practical assumptions explicit. For instance, in the case of semi-autonomous vehicles, understanding how the human driver adapts to the presence of different automatic driving assists helps to design safer and efficient driving. Moreover, models can be used to generate predictions of how drivers adapt to various design choices. However, modelling is time-consuming and hard, and involves a learning curve that may be unacceptable. In order to facilitate proliferation of computational cognitive modelling, the modelling workflow must be made usable. This involves creating tools for modelling, but also benchmarks for model testing, and "canonical scenarios" for trying out design and modelling ideas in familiar environments.

3.17   Much more data is needed

Wendy Ju (Cornell Tech – New York, US)

To build computational models for human-AV interaction, we need a lot more data. Of course, we have all heard that "all models are wrong, some models are useful." To date the models I have seen in the space of human-AV interaction do not account for the cultural variation, contextual variation and temporal variation that we know for certain would affect any predictions for how people will respond to an AV.

Instrumentation

One issue with gathering data to address cultural and contextual variation, at least, is that is is unlikely that any one research group in one location would be able to gather data across two sites to even begin to address this issue. For this reason, the question of publishing and sharing the way that instrumentation occurs in lab and real-world data collection. Sharing instrumentation set-ups makes it easier to have datasets that are comparable; for this reason, publications that discuss and benchmark different instrumentation configurations for data collection of human-vehicle interaction should be considered in and of themselves to be contributions to the community.

Scale

The scale of the data that would need to be collected to address cultural, contextual and temporal variation is also an issue. Finding ways to use machine learning to augment and scale human observation and coding in empirical data analysis is critical. As a community, we should be discussing computational methods to share scale and validate data analysis.

Data

I believe that part of what is needed is that we as a community need to treat the gathering and sharing of data as a real contribution, even before it has been analyzed or yielded any insights. Empirical data, particularly from real-world studies, is often "dirty"– equipment failures, participant strangeness, and exceptional events abound. Given the sparsity of of this much needed data at all, it is more important that any issues with the data be clearly documented and explained than that we demand great perfect data that fits our models.

3.18 Computational cognitive models for attention monitoring

Tuomo Kujala (University of Jyväskylä, FI)

Because of the complex and unpredictable nature of traffic with human road users, I argue that in order to be safe, driver should monitor the behavior of the car and the traffic environment up to SAE level 3 of automated driving (maybe even at level 4). This requirement is decreased as the automation level increases and that is actually a big problem as the driver's attentional capacity is freed but not fully. SAE levels are also somewhat problematic in this sense as a real car given a level 3 status may not always function as a level 3 car should (e.g., always know when the driver should be alerted to take over).

Current driver attention monitoring systems are really impressive in their object and behavior classification performance but I would argue that it is not enough to monitor the driver or insides of the cabin to know if the driver is inattentive towards driving. Drowsiness is probably the only form of inattention that can be reliably detected by monitoring the driver only. The attentional demands of driving vary based on traffic situation, surroundings, upcoming situation, etc. and there is often so-called spare attentional capacity in driving. And even more when the level of automation increases. But how much, that is the question? Definition of these requirements might also get harder when we go up in the level of automation. As this task is fairly complex, we would need computational models for this.

For these reasons, and because driving is a safety-critical context, we'll need also prescriptive computational models, not only descriptive and predictive. For instance, we would need a normative criterion to define when the driver is inattentive, depending on situational and driver-specific variables

3.19 In-vehicle human-automation interaction: Bumper-to-bumper traffic, and short bursts of activity in non-driving tasks

Andrew Kun (University of New Hampshire – Durham, US)

A driving scenario that promises to be important for human-automation interaction in vehicles in the near term is bumper-to-bumper traffic. In this scenario the vehicle is relatively easy to control by automation, and there is a relatively low risk of injury in case the automation makes a mistake. At the same time there would be a large positive impact on the driver if they could use the time that the vehicle is in bumper-to-bumper traffic to engage in non-driving tasks.

How can drivers use their time if the car can drive itself in bumper-to-bumper traffic? The answer is: short bursts of non-driving activity with rapid returns to driving. The reason is that bumper-to-bumper traffic doesn't last forever – it's likely to often take only minutes. And transitions back to driving will need to be very quick, on the order of seconds, not minutes [1].

It is likely that those short bursts of activity will very often involve manual-visual tasks. We recently conducted a time-use study with 400 knowledge workers who commute by driving [2]. We asked them what they would like to do in a future, safe automated vehicle. They provided us with tasks they are interested in, and we assessed the tasks in terms of the need for various cognitive resources (using the Wickens multiple resources model). We found that our participants wanted to do more manual-visual tasks, both for work and for personal tasks. Thus, we expect more typing and browsing, in contrast to silent reflection or listening to music.

References
1 Nagaraju, Divyabharathi, Alberta Ansah, Nabil Al Nahin Ch, Caitlin Mills, Christian P. Janssen, Orit Shaer, and Andrew L. Kun. (2021) How Will Drivers Take Back Control in Automated Vehicles? A Driving Simulator Test of an Interleaving Framework. In *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 20-27).
2 Teodorovicz, Thomaz, Andrew L. Kun, Raffaella Sadun, and Orit Shaer. (2022) Multitasking while Driving: A Time Use Study of Commuting Knowledge Workers to Assess Current and Future Uses. *International Journal of Human-Computer Studies*.

3.20 Four challenges to computational models in Human-AV interaction

Dietrich Manstetten (Robert Bosch GmbH – Stuttgart, DE)

Computational models (including but not limited to cognitive models) are playing an essential role to enable safe automated driving and to make it a reality on tomorrow's roads. We are looking at four challenges in the field.

1) **Modeling driver engagement and driver availabilty.** Driver monitoring becomes a must in the homologation of assisted and automated driving. We need real-time models estimating whether the driver fulfills his supervision task in Level 2 (engagement) and whether he will be able to take over if a request occurs in Level 3 (availability).
2) **Take-over performance models.** Respecting the current status of traffic, vehicle, and driver we need models to predict the temporary and quality aspects of the take-over activity and to decide if his/her actions are intentional and supportive for safe driving.
3) **Shared control / cooperative driving.** As humans are still needed for subtasks in automated driving (e.g. for deciding on lane changes and other tactical aspects)driver's computational models should be used during system development to analyze and validate the successful cooperation between driver's actions and automation control.
4) **Good driving behavior as a role model for automation.** Behavioral models describing the human driving process in traffic are a good way to serve as a role model and to be mimicked by future automated and even autonomous driving systems. Separating "good" and "bad" driving can help that autonomous driving will be able to realize a chauffeur-like safe and comfortable driving style.

3.21 Computational models of humans as tools for enabling safe and acceptable vehicle automation

Gustav Markkula (University of Leeds, GB)

Computational models of human cognition and behaviour have applied uses in, and in some cases will be indispensable for, development of safe and human-acceptable vehicle automation. At lower levels of automation (e.g., SAE level 2-3), models of how humans monitor the automation and take, receive, or share the vehicle control can be used (1) as part of online, real-time algorithms to infer driver state and predict driver actions, and to adapt vehicle behaviour accordingly, or (2) in offline computer simulations to evaluate different automation design alternatives. The same division into use in both online algorithms and offline testing applies also for higher levels of automation (e.g., SAE level 4-5), but then with respect to modelling the cognition of behaviour of other road users around the automated vehicle. One exciting prospect and grand challenge, across all of those application areas, is to find the balance and combinations between mechanistic/cognitive models on the one hand, and data-driven/machine-learned models on the other, to enable extensible models that users (automated vehicle designers/engineers) can generalise to new contexts and data sets, without direct involvement from the original model developers.

3.22 Data for and models on Trained Operators

Nikolas Martelaro (Carnegie Mellon University – Pittsburgh, US)

Current thinking around human-machine interaction in autonomous vehicles is often through the lens of untrained everyday users. We may often think that our models of human behavior are generalizable. However, after training, people may develop new ways of acting, especially

with autonomous systems. ,A potential area of work that is underexplored in the space of autonomous vehicles is how trained operators will act. For example, a bus operator or a semi-truck driver will have more training in both driving and with the systems they are using. As such, cognitive models for these people may be different than those of a lay population. The research community should consider research on trained operators and should capture data in experimental settings and develop models that are specific for them. Such work may lead to better predictive models that can be used in real-world autonomous vehicle sooner than in vehicles with a lay population. This line of research may also identify good mental models that could then be taught to lay drivers to help improve their interactions with autonomous vehicles.

3.23   Simulation Intelligence and Computational Models in Autonomous Vehicles

Roderick Murray-Smith (University of Glasgow, GB)

"Simulation Intelligence" brings together first principles models, probabilistic programming and machine learning. It has a lot of potential for supporting a more principled approach to creation of computational models, which can help support and direct scientific progress in the field. Moving from correlational fits to data towards causal models which really represent the actual behaviour will be vital for reliable behaviour in novel contexts. Using forward and inverse inference mechanisms, as is common in other areas of science, gives a potentially more reliable way of formulating the problems, and managing models for components of larger models. Closed-loop aspects are also critical – they affect how we acquire data, and also provide challenges associated with the interdependence of human behaviour on the dynamics of the vehicle (e.g. crossover models), but this seems to be underconsidered at the moment.

3.24   Computational Rationality as an Emerging Approach to Inform the Design of Interactive System

Antti Oulasvirta (Aalto University, FI)

How do people interact with computers? This fundamental question was asked by [1] with a proposition to frame it as a question about human cognition, in other words as a question of how information is processed in the mind. Recently, the question has been reframed as a question of adaptation: how do people adapt interaction to the limits imposed by cognition, device design and its environment? The core assumption of computational rationality is that users act according to what is best for them given the limits imposed by their cognitive architecture and their experience of the task environment. The theory can be expressed in computational models which explain and predict interaction and be therefore used for design and adaptive systems.

References

1   Card, S. K., Moran, T., & Newell, A. (1983). The psychology of human-computer interaction. Hillsdale, NJ: L. Erlbaum Associates Inc.

3.25   A common understanding of models...

Andreas Riener (TH Ingolstadt, DE)

In many discussions I have had over the last years with my students (in the interdisciplinary program User Experience Design), within my research group but also with colleagues (e.g. in the context of Dagstuhl Seminar 16262), it has turned out again and again that there is a different understanding of what a model is and how or for what it can be used. The main reason was that we have different backgrounds (engineering, computer science, design, social sciences, psychology). We need to find ways and means to make sure we ("the system designer") mean the same thing... Predictive models could be a common basis – but of course they are only a simplification of real behavior [1]. If we focus on human-machine (or human-AV interaction), it is particularly necessary that both parties (human, machine) understand each other and can mutually assess how far the partner is trustworthy in a certain situation [3]... And this is particularly important in the recent research field "Cooperative Driving" – where automation and human work together as team players [2]. In a team, each agent must be informed about the other agents' current activities, their strategies, the status of their efforts, if they are having problems, and their intentions for planned actions. The Object, View, and Interaction Design (OVID) model [1] (Roberts, 1998) contains three design models based on Don Norman's notion of cognitive engineering [4] and clearly visualizes the problem: We cannot really design a user's conceptual model, as this is highly inter-personal (but also intra-personal), based on experience, and exist only in the brain of a person... Furthermore, user interface and vehicle designers as well as user experience practitioners are often challenged with the question for which user group they are designing for – each with different needs, different interests, and very different ways of interacting with technology. In vehicle production, we do not have the luxury of focusing on only one group (at least not so far), i. e., designing vehicles for specific age groups or cultures, that's why interaction designers and engineers must learn to recognize and reconcile the needs of their main user demographics. This problem will remain even with fully automated driving when using the car as a place for relaxation, entertainment or work. Defining characteristics, differences, and tensions between individual user groups might help to account for different individuals. An important question in this regard is, whether or not there is a single system suitable for all (or at least most) customers and stakeholders. Is it axiomatic to target user groups differently? (HMI and system configuration allow to...). Dagstuhl Seminar 22102 was the perfect place to discuss with all the participants from different backgrounds (my access is hypotheses-driven experimental research) the possibilities of combining existing models, extending/improving them, or developing new models to better represent the wide variety of variants in human-AV interaction/cooperation and human diversity. I hope that results can be derived from this week's collaboration that will help the community....

References

1 William Hudson (2001).Toward Unified Models in User-Centered and Object-Oriented Design. In *Object Modeling and User Interface Design Designing Interactive Systems*. Addison-Wesley, Pages: 313-362. ISBN: 0201657899 (Figure 9.4; p. 328).

2 Frank Flemisch, Matthias Heesen, Johann Kelsch, Julian Schindler, Carsten Preusche, and Joerg Dittrich (2010). Shared and cooperative movement control of intelligent technical systems: Sketch of the design space of haptic-multimodal coupling between operator, co-automation, base system and environment, Vol. 1.

3 Philipp Wintersberger, Anna-Katharina Frison, Andreas Riener, and Linda Ng Boyle (2016). Towards a personalized trust model for highly automated driving. *Mensch und Computer*, Workshopband, 2016.

4 Norman, D. A. (1986). Cognitive engineering. In D. A. Norman and S. W. Draper (Eds.) *User Centered System Design* (pp. 31–61). Hillsdale NJ: Erlbaum.

5 Martin R. Baumann and Josef F. Krems (2009). A Comprehension Based Cognitive Model of Situation Awareness. In V. D. Duffy (Ed.), *Digital Human Modeling*, Vol. 5620, pp. 192–201, Springer, https://doi.org/10.1007/978-3-642-02809-0_21.

3.26 Anticipating the cognitive state of the driver

Nele Rußwinkel (TU Berlin, DE)

The next big challenge for computational models of human-Automated Vehicle Interaction is, from my perspective, to integrate more aspects and understanding of the human perspective. I do not think that one modelling method needs to capture all possible aspects of a scenario and related entities to be a good model. What is needed are good models that can capture specific aspects and a good concept that can integrate this information with emerging transparency – how different aspects have led to a specific understanding of the situations. A modular approach might be helpful to achieve this in case some module results have a symbolic form. What is needed is a tighter connection of human and technical system understanding. It is not sufficient to measure some mental states, it is also (in some situations) necessary to understand what the cause of the measured state is and how to address it. I claim for more research on models that anticipate the human driver in the dynamic situation and include mental representations e.g., for "Situation Awareness" or "Mental models" or specific "Expectations" about what the human considered relevant now and what is expected to happen next. This would enable a better mutual understanding and a more natural (less effortful) shared control of the car or the situation and to avoid misunderstanding.

3.27 Models that Inform Design

Shadan Sadeghian Borojeni (Universität Siegen, DE)

The advances in automation and machine learning are changing the designs of in-vehicle user interfaces from static command-response "buttons & displays" to dynamic adaptive interfaces that adjust themselves to user behavior and preferences. From a modeling perspective,

however, this can be a challenge to integrate these new "emerging behaviors" into the models. On one side, having simplistic models allow us to scale them to different scenarios and interaction forms, on the other side, it restricts our ability to precisely describe/predict the adaptive and dynamic interactions between the user and the automated vehicle.

Another challenge is answering the question of whom we are designing for and in which scenario? The first step towards that is defining the sides of the interaction, i.e., the human and the automated vehicle. Identifying the capabilities and responsibilities of each of these partners across different levels of automation can help us understand what factors are required to be considered in the models. For example, for an urgent takeover situation in level 3, factors that define users' situational awareness, the urgency of the situation, and contextual factors might be determinants of the behavior. However, for a level 5 scenario, where the interaction goal is mostly to assure users' comfort and wellbeing, the determining factors stem from users' psychological needs and affective state rather than her situational awareness. It is, therefore, necessary to first define users' tasks at each level of automation and in different scenarios. And then elicit the factors that are required to fulfill these tasks' goals. This allows us to create a modular model of tasks, levels of automation, and required factors that assist us to identify the important elements that should be addressed in the design of in-vehicle user interfaces.

3.28   Complexity of computational models

Hatice Sahin (Universität Oldenburg, DE)

Computational models can help to ease the cost of experimental studies. Discussions have been developed under the consensus that various individual or external factors affect distribution of driving responsibility between the human and the automated vehicle. These factors range from personal preferences, trust, experience, social and legal norms to environmental conditions. While it is crucial that these factors are recognized, it may not be vital for computational models to include all possible parameters. It is still a matter of discussion whether their benefit is limited by their simplistic qualities. While some might consider that they might have to be complex enough to provide generalizability, they could still be helpful to investigate individual factors with limited complexity.

3.29   Teamwork and Driver-Vehicle Cooperation

Boris van Waterschoot (Rijkswaterstaat – Utrecht, NL)

Main reference Sebastianus Martinus Petermeijer, Angelica M. Tinga, Reinier Jansen, Antoine de Reus, Boris van
Waterschoot: "What Makes a Good Team? – Towards the Assessment of Driver-Vehicle
Cooperation", in Proc. of the AutomotiveUI '21: 13th International Conference on Automotive User
Interfaces and Interactive Vehicular Applications, Virtual Event / Leeds, United Kingdom,
September 9-14, 2021, pp. 99–108, ACM, 2021.
URL https://doi.org/10.1145/3409118.3475153

With the introduction of driving automation systems [1], the subtasks of the DDT are being performed by the (human) driver, by the driving automation system, or by both. This means that under certain circumstances, human and automation have become collaborative partners in executing (parts of) the driving task.

Although the view of driver and automation acting as collaborative partners has been well established, means to assess their teamwork are still lacking. Moreover, available evaluations are usually addressed either from a technical stance or from a human factors viewpoint, which does not comply with a general acknowledged view of a unified driver-vehicle system. This stance on teamwork and a description of our aim to evaluate driver and vehicle cooperation by means of a framework can be found in [2] .

Our work is currently being followed up by a roadmap – which is shared as soon as possible – aiming to deliver guidelines on the framework's practical implementation, potentially supporting monitoring, evaluation and design activities concerning driver-vehicle cooperation.

This abstract is a call to get involved in our activities regarding the objective evaluation of driver-automation cooperation.

References
1   SAE, Society of Automotive Engineers. (2018) Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles. J3016 JUN2018. SAE International: Warrendale, PA, USA.
2   Petermeijer, S. M., Tinga, A., Jansen, R., de Reus, A., & van Waterschoot, B. (2021). What Makes a Good Team? Towards the Assessment of Driver-Vehicle Cooperation. In *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 99-108).

3.30   From Trust Studies to Trust Models

Philipp Wintersberger (TU Wien, AT)

Six years have passed since my first Dagstuhl Seminar. At that time, the first deadly crash with an automated vehicle happened. It became clear that the trust relationship between drivers and automated vehicles will become a key issue for guaranteeing safety in the future since drivers are not necessarily automation experts. Thus, we intensively discussed the topic of trust in automated vehicles and proposed to research "personalized trust models" so that in-vehicle HMIs can adapt to drivers' states, strengths, and weaknesses. A great variety of trust studies have been conducted and published since that seminar, revealing many relevant factors influencing this multidimensional psychological construct. However, a sophisticated trust model has still not been developed. Thus, I am very happy for the invitation to this modeling seminar, as I got the chance to (1) talk to modeling experts from my domain and (2) discuss with colleagues concrete, actionable steps to move towards the goal outlined above. I am confident that it will not take another six years to integrate all the past study results and develop an initial functional version of the proposed model.

3.31 Understanding and challenges of computional models of human-automated vehicles interaction

Fei Yan (Universität Ulm, DE)

In order to better understand the interaction process between humans and automated vehicles as well as develop adaptive assistance systems to support drivers, computational models that can precisely predict driver behaviors are needed. However, with the change of automation levels, there are challenges while developing computational models. With the change of driver's role from SAE L2 to SAE L3, the focus of modeling changes from monitoring driver's states, to modeling driver's takeover reactions which are more safety critical compared to SAE L2. From SAE 4 or SAE L5, the transition of responsibility can be relevant. While modeling takeover process from SAE L3,indicators like takeover timing as well as takeover quality such as maximal acceleration and TTC, subjective measurement such as workload can be taken into account. Facing the complex scenario in the mixed traffic, the combination of cognitive approach with black-box models is needed, which can maximize the advantages of different models, but also ecologically save the cost of developing complex cognitive models. The cognitive theory, evidence or empirical research should be made use of to help to develop casual black-box models, which can help to interpret the relations between relevant factors.

4 Working groups

4.1 Learning and Adaptation

Jelmer Borst (University of Groningen, NL), Alexandra Bremers (Cornell Tech – New York, US), Birsen Donmez (University of Toronto, CA), Mark Eilers (Humatects – Oldenburg, DE), and Roderick Murray-Smith (University of Glasgow, GB)

Learning was discussed from two points of view: on the one hand users have to adapt to new (autonomous) systems, and on the other hand models need to be able to learn and adapt.

With respect to users, they have to learn operating (semi-)autonomous vehicles, and also adapt to such vehicles being present in the environment. One challenge is that different systems (e.g., Tesla, GM) might behave differently, which makes learning harder. In addition, it is often not clear whether they operate in autonomous mode or not. Second, people might "abuse" the new systems, for example crossing the road directly in front of an autonomous car, as it will stop anyway.

Models will also have to adapt to new situations on the road (i.e. more autonomous cars) and to adjusted human behavior. We can differentiate between quantitative changes (new speed limit) and qualitative changes (new traffic rules). While the former is probably possible, the latter is much more difficult to automatically adapt to. In general, we expect that adaptation, in particular qualitative, is easier for cognitive and causal models, as these take content into account.

To adapt machine-learning-based models, one could think of using feedback of the user. However, some drivers might be more amenable to give feedback than others. We therefore recommend implicit learning, based on the behavior or physiological responses of the users.

4.2 How can model use be increased in design?

Alexandra Bremers (Cornell Tech – New York, US), Jussi Jokinen (University of Jyväskylä, FI), and Gustav Markkula (University of Leeds, GB)

There are examples of models used by designers and engineers in the industry, such as for traffic testing. However, a question remains about how accessibility can be improved, and adaptation of models can be increased.

Potential roadblocks and limitations for using models in design are:

1. Cost-benefit tradeoff. An investigation is needed to clearly map out scenarios where modelling adds the most value in addition to testing methods.
2. Design space restriction due to specificity of models. For instance, a typing model would assume touch-based interaction within a specific limited interface size.
3. Language intersection between the designer's creativity and the model's formalized way of using specifications. In machine learning, tools have been developed, such as the OpenAI gym, which bridges the gap between model and application.
4. Complexity depending on the task. Tasks that seem simple from a design perspective, such as selecting the best option in a webshop, could involve a very complex cognitive model.

In AV applications, HMI and UI design seem ripe for model integration. Successful tools exist, such as Distract-R and OpenAI Gym, online design tools such as the Adobe Mixamo library, and wireframing tools like Miro and Sigma. Combining these could result in a tool for both quick model-assisted sketching and a more thorough model-based evaluation. The behaviour of other road users, such as children playing soccer on the sidewalk, could be a lot more complex to realize. Discussions with AV industry engineers could further confirm specific needs and requirements.

4.3 How can empirical research support computational models (and vice-versa)?

Debargha Dey (TU Eindhoven, NL), Moritz Held (Universität Oldenburg, DE), and Fei Yan (Universität Ulm, DE)

In this focus group, we dove deep into the caveat of empirical research and computational modeling approaches and discussed the situations where each is applicable, and where one approach can benefit from the other. While a modeling approach is not needed to answer every research question, it can be critical if we are trying to justify cognitive processes (white box models), or trying to reduce the complexity of the research space by looking at causality and/ or relationship between certain parameters of the environment and the observed behavior (black box models). We recognized that there can be no one strict guideline to the question of the data points needed to develop a model (as it largely depends on the research question). In situations when it is difficult to estimate the impact or relevance of certain parameters for heuristic approaches, modeling approaches can help by highlighting

them. Thinking through the cognitive architecture of a white-box model can identify the "independent variables of interest" for empirical studies. Furthermore, data from empirical research can be helpful in building hypothesis for computational models especially predictive models using machine learning techniques. However, "absolute" black box models lack the transparency and explainability in terms of causality explained through cognitive processes, which calls for inputs from empirical research and cognitive theory towards a combined approach that leverage the best of each world.

4.4 Modeling Long-term Effects in Human Technology Engagement

Patrick Ebel (Universität Köln, DE), Hatice Sahin (Universität Oldenburg, DE), and Philipp Wintersberger (TU Wien, AT)

We discussed how long-term effects (distribution shifts) in human behavior when interacting with technology can be modeled.

To make the problem more tangible, we have decided on a specific use case: "How to model the probability for a driver to activate automated driving functions?" Such an approach could be an interesting element for driver monitoring systems, since usage patterns can be related to interaction problems (such as trust) [1]. Given a certain set of input variables (environmental context, individual driver characteristics, trust in the system, effect of "events", etc.) the model should predict how likely it is for the driver to activate driving automation. The goal of such an approach is to (1) model long-term driver behavior and to (2) develop effective measures to increase the usage of automated driving functions.

To model the probability of human engagement with technology, three main components are be modeled over time:

1. The general probability of engagement. This probability is dependent on factors like the benefit of using the technology (How good is the driving automation?), learnability (How well do drivers adapt?), trust (Do Drivers trust automation?), individual characteristics, and the probability of being in a situation where the technology is applicable (is it possible to activate automation in a specific driving situation?). This probability can be modeled as a continuous function over time and serves as the base probability of engagement.
2. The influence of incidents. We defined incidents as event-based disturbances that influence the overall probability of engagement at a certain point in time. They can be modeled using a step function. An example would be the sudden change of the activation probability after experiencing a critical situation while driving with automation activated. A relevant question in this regard is, if the underlying cause for a behavior change can be derived from a changing usage pattern (i.e., do there exist detectable patterns for certain events)?
3. The effect of interventions. An HMI intervention is provided to a human to mitigate the effect of an incident. Interventions can be modeled by adding an "intervention effect" function to the general engagement function. The idea is that a positive intervention like explaining why a critical situation couldn't be solved by the automated driving function can change the probability of engagement over time such that it approaches the original probability before an incident happened more quickly.

Subsequently, we developed a first sketch for a study to investigate the three main components of the theoretical model. This study needs to be further developed.

References

1   Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. Human factors, 39(2), 230-253.

4.5   Focused session "Application in specific scenarios"

Martin Fränzle (Universität Oldenburg, DE), Luisa Heinrich (Universität Ulm, DE), and Roderick Murray-Smith (University of Glasgow, GB)

As applications and application scenarios make their way into the research stream mostly via use cases and benchmark examples, the discussion in the group focused on how to obtain representative domain coverage via benchmarks and use cases. The lack of a reasonably broad and generally accepted set of such benchmarks and use cases was identified as an impediment both to rigorous and comparative assessment of the current state of research and to the identification and prioritization of open research problems. An informed, problem-driven advancement of the state of the art would thus require to find, collect, and publicize benchmarks that collectively obtain representative coverage of the relevant applications and application scenarios and that facilitate rigorous evaluation of and friendly competition between modelling approaches.

This quest in turn induces the need for an intelligible characterization of benchmarks that permits a mapping of individual benchmarks as well as benchmark sets concerning their particular contributions, thereby supporting demand-driven selection and adoption of existing benchmarks. For this mapping of the area, we came up with the following dimensions and explications of dimensions:

1. The dimension of paradigm of use of the human models, distinguishing between human models as proxies of human behavior in model-based design and the embedding of human models into an operational system.
2. Means of parametrisation and validation, namely observed in real-life / on real road / by physical test track driving or through simulator studies.
3. Focus of the embedded human model, ranging over driver only (possibly including in-car passengers also), other traffic participants (in alter cars, in environment – especially VRUs), or both, including reflection of mutual reactive behavior.
4. A broad set of quality criteria for the benchmark specification, including among others heterogeneity of driver behavior and coverage of driver types as well as societal groups (e.g., elderly), non-discrimination, explicit identification of observable/measurable correlates for latent or hidden phenomena of interest, strong correlation of benchmark-represented/benchmark-measurable features to real-world effects and effect strengths, accurate coverage of variability, not just of nominal behavior, as well as means for automated variant generation to avoid overfitting to particular benchmarks, test-retest reliability, inter-rater reliability, and repeatability and reproducibility.
5. Aim of benchmark, covering a wide range of attributes from incentivizing research through friendly competition to facilitation of rigorous relative evaluation and ensuring coherency of methods, quality criteria, etc. between scientific communities. The last dimension identified was 6. technical prerequisites like executability on common hard- and software platforms, convenient packaging (e.g., via containers), and issues of open access and open science.

The discussion group holds the belief that such a classification would enable consolidated and collaborative research, stabilize and align the pertinent quality criteria across domain-relevant disciplines, and thereby advance scientific progress. Necessary first steps to render it reality would be to, first, generate a questionnaire permitting individuals and research groups to locate their respective benchmarks in the above multidimensional criteria set, second, to establish a platform for publication of benchmarks, including video tutorials and supportive contacts, and, third, to foster industrial contribution.

4.6 Towards common tests for Models

Christian P. Janssen (Utrecht University, NL), Jelmer Borst (University of Groningen, NL), Benjamin Cowan (University College – Dublin, IE), Justin Edwards (ADAPT Centre – Dublin, IE), Mark Eilers (Humatects – Oldenburg, DE), and Jussi Jokinen (University of Jyväskylä, FI)

We discussed whether there can be a benchmark test, or evaluation method, to better compare different models. We were inspired by how competitions are done in the language community. In that community, there are multiple competitions and test frameworks. Not any of those is the "killer test", but across tests one can see where a model or framework is stronger or relatively weaker. This moves the community forward.

In the field of computational modeling for human-automated vehicle interaction there are perhaps two challenges:
1. What is a good model that can perform well on an evaluation test (or tests)?
2. What is an appropriate (large enough, diverse) dataset to test these models on?

We foresee that this can be approached in three steps, that each require attention:
1. Identify what the possible criteria are to evaluate a model on and define what scenarios are needed to test these.
2. Collect data in an empirical scenario where the above features emerge
3. Develop models that can then be tested on this dataset / scenario.

In later discussions we delved deeper into question 1. As a constraint we set that the model should exhibit human-like behavior. We looked at other papers that have defined model criteria, in particular papers by Anderson and Lebiere [1] and Taatgen and Anderson [2]. For these papers, we went through the criteria they defined and determined if and how they apply to the driving domain. We've identified that some modifications and specifications are needed. We also foresee that additional criteria – such as the ability to perform driving tasks and driving specific tests in a human way – need to be added. We plan to further refine these ideas in a follow-up project.

References
1 Anderson, J. R., & Lebiere, C. (2003). The Newell test for a theory of cognition. *Behavioral and brain Sciences*, 26(5), 587-601.
2 Taatgen, N., & Anderson, J. R. (2010). The past, present, and future of cognitive architectures. *Topics in Cognitive Science*, 2(4), 693-704.

4.7 What are important scenarios?

Dietrich Manstetten (Robert Bosch GmbH – Stuttgart, DE), Moritz Held (Universität Oldenburg, DE), Nele Rußwinkel (TU Berlin, DE), and Fei Yan (Universität Ulm, DE)

In the beginning of the discussion, it became quite obvious, that we can't write down all important and relevant scenarios. We searched for a framework-like setup, where a scenario is composed of a multitude of independent factors. Some sort of building blocks to construct a specific scenario.

Several of these factors are influenced by the environment and the actual situation. They are relevant, but not directly linked to the automation. We identified

- Traffic environment. Freeway (number of lanes, exit section, tunnel, construction site) / Rural road (curvature) / Urban traffic (crossing, pedestrian, cyclist, parked vehicles)
- Criticality of situation (mostly described by dynamic behavior of others)
- Vehicle Type (truck, passenger car)
- Driver characteristics (stable factors as age / driving experience / knowledge about automation, dynamic factors as time pressure / fitness level)
- Boundary conditions (weather, daylight/night, road surface)

When it comes to human-AV interaction, the main ingredient of the scenario comes from the automation itself. There is the temporary aspect of the current phase of automation, which can pose specific questions: activating the automation, during automation, transfer-of-control. Within these phases, again we see multiple factors, and we partly need to separate carefully between the levels of automation, as some of the effects are relevant for a specific level only.

- Level 2 asks for continuous driver engagement and anticipatory behavior. The transfer of control is usually initiated by the driver.
- Level 3 needs driver availability during automation. Fast set-up of situational awareness after a take-over request. Assessment by take-over time & quality (including safeguard, e.g. mirror glances).
- Level 4 changes the time scale more to minutes than seconds. Some requirement of perceptibility. On system level, there may even be the need for a driver lock-out avoiding unintentional or undesired actions.
- For level 3 and 4 reaching the ODD limit or possible system failures becomes most relevant situation.
- Information design and communication on all levels (dependent of specific situation)
- Clear definition of visual behavior requirements (level 2), cognitive requirements (level 2 and level 3)
- Mode shift with transfer between levels (e.g., from level 3 to level 2) and corresponding challenges of mode awareness
- Aspects of trust towards the system.

The presentation of the group's work in the plenum added the question, where models can be part of the system. This augments the perspective, that knowledge represented in driver models comes partly as a requirement from legal perspectives and regulations. Consequently, the constraints from a legal perspective have to be respected carefully.

4.8   What is needed for different SAE level?

Dietrich Manstetten (Robert Bosch GmbH – Stuttgart, DE), Martin Baumann (Universität Ulm, DE), and Shadan Sadeghian Borojeni (Universität Siegen, DE)

In this focused session we discussed, which specific system design possibilities are offered within the different SAE levels of automation. We came out with a sketch of a framework to clearly describe the user tasks in the different levels. This description has the potential to serve as a backbone for the system design.

With respect to the levels, we concentrate on the levels 2, 3 and 4. By this, we are adapting somewhat the terminology used by the German Federal Highway Research Institute BASt of assisted (level 2 and below), automated (level 3) and autonomous (level 4 and above) driving, see `https://www.bast.de/EN/Automotive_Engineering/Subjects/f4-user-communication.html`. The time scale for the user's tasks varies as well within these levels, from below seconds (level 2) via few seconds to minute (level 3) to minutes (level 4).

During the automation phase, the main task is shaped by monitoring. Monitoring means to monitor the environment, to monitor the system, and to monitor yourself. The monitoring serves the goal to decide whether I (as the driver) need to intervene, which would end the monitoring phase, and replace it by an action.

We systematically added the view for the phase transitioning from on to off. The possible interventions can be differentiated as an adjustment of the automation behavior, a system-initiated taking back of the control (as a result of a request-to-intervene), and a driver-initiated taking back of the control (might be necessary for safety reasons or just a current driver's preference). As during the automation phase, depending on the level of automation, the concrete task description varies again. We are explicitly looking at the "what?", "why?" and "when?" of the transition.

The two main pictures compiled by the working group are added to this report. The group had the clear impression to have a solid ground for an overall framework, but being clearly away from being complete. The focus was so far from the user's perspective only, the system's perspective needs to be added. The group plans to pick it up at a later stage, and to compile a full view out of it. Replacing detailed settings in scenarios by a comprehensive task-oriented view. To be continued.

4.9   Multi-agent modelling: Platooning use case

Gustav Markkula (University of Leeds, GB), Lewis Chuang (LMU München, DE), Debargha Dey (TU Eindhoven, NL), Patrick Ebel (Universität Köln, DE), and Christian P. Janssen (Utrecht University, NL)

In this focused session, we discussed a number of different scenarios involving multiple human road users that could benefit from the application of human behaviour models. Our discussion converged towards a scenario, centered on platooning – wherein a number of vehicles (cars and/or trucks) follow each other on a highway, at close distance in automated mode, typically

Driver tasks according to SAE levels for the transition phase "on to off"
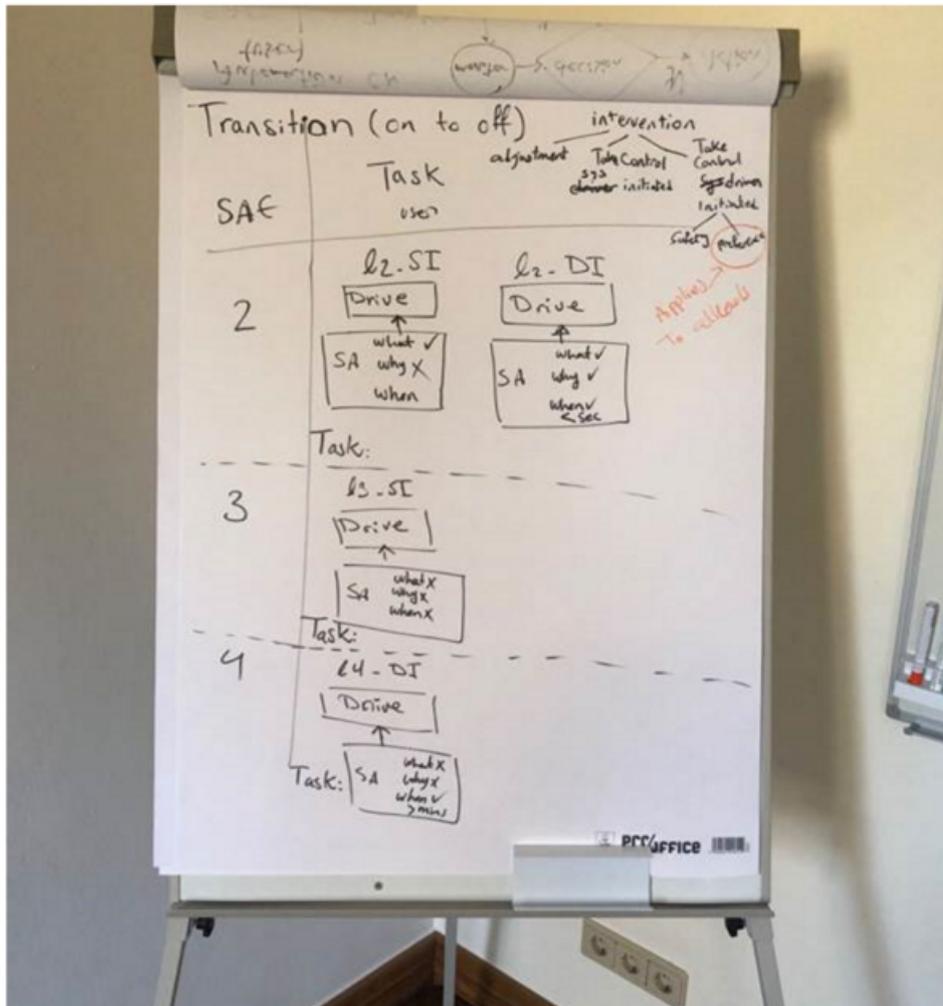
with the expectation that drivers are available to resume vehicular control upon demand. This rich setting gave rise to a number of different control transition scenarios. In particular, we discussed the requirements for enabling the platoon to temporarily break apart to change lanes – e.g., due to an upcoming lane closure – and reassemble later. After discussing the potential contributions of different types of human behaviour models, we agreed on three main types of uses: First, models of the drivers in the lead and following vehicles can inform online algorithms in the platooning vehicles. This would guide real-time decisions how to break up the platoon, by predicting likely human responses to different alternative vehicle actions, and by choosing the control approach that is predicted to give the best outcome in terms of safety, efficiency, and comfort. Second, similar (but possibly not identical, since they would serve slightly different purposes) driver behaviour models could be useful in offline simulations. These would assist an engineer in designing systems by providing possible platoon breakup scenarios for automation implementations and testing – for example, to test lead time requirements for breaking up the platoon. Such tests would optimize safety and comfort, as a function of platoon speed, traffic density, etc. Third, we discussed that

Driver tasks according to SAE levels for the automation phase
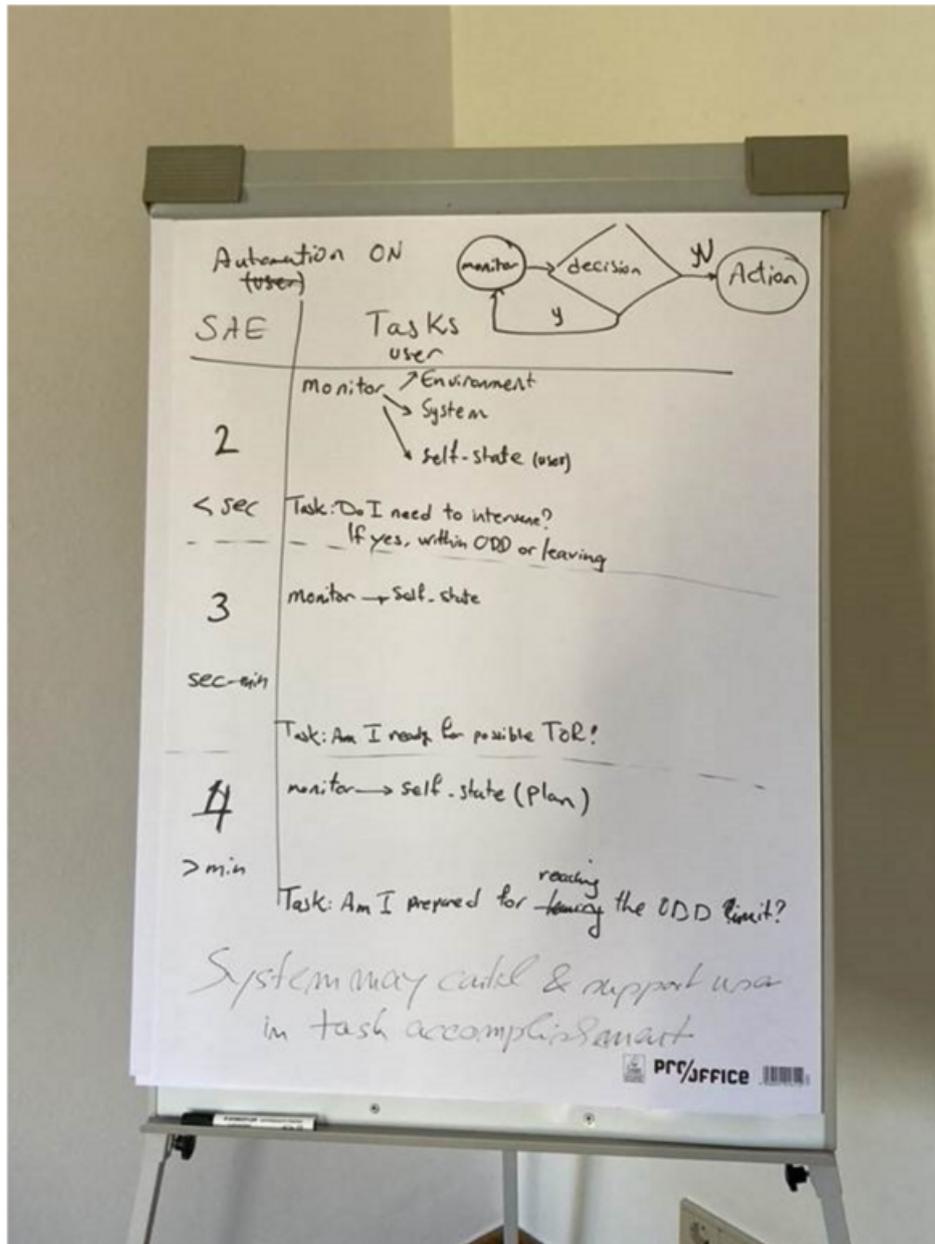
models could be useful on a conceptual level. Such models would help system designers and decision-makers better understand the behaviour and limitations of human drivers across different situations.

4.10 What computational models are available for human-automation interaction?

Gustav Markkula (University of Leeds, GB), Debargha Dey (TU Eindhoven, NL), Moritz Held (Universität Oldenburg, DE), Nele Rußwinkel (TU Berlin, DE), and Fei Yan (Universität Ulm, DE)

In this focused discussion we attempted to provide as broad as possible an overview of currently existing computational models which have been applied to the study of human interaction with automated vehicles, or could be applied in such a context. We listed a large number of references (which will be included in a separate document by the seminar organisers), and as we did so we structured the existing work in a rough taxonomy including the following dimensions: (1) Type of traffic scenario addressed, e.g., automation control transitions, shared control, multi-agent interaction, teleoperation, or driver support systems. (2) Type of (intended) use of the model, e.g., use of models to conceptually guide automation design, use in online algorithms for prediction or decision-making as part of the automation itself, use in offline simulated testing of automation, or as human safety benchmarks for automation. (3) Aspect(s) of human cognition/behaviour modelled, e.g., vehicle control, visual behaviour, individual characteristics, or unobservable mental/cognitive states. (4) Type of model, e.g., machine-learned, cognitive architecture, control-theoretic, and models focused on decision-making. These dimensions are to a large extent independent, but we also discussed that especially different model types, not least distinguishing between machine-learned, black-box models and mechanistic, white box models, might map more or less naturally to different parts of the other dimensions in this taxonomy. We discussed the strengths of these different types of models with respect to time critical decisions, understanding, drivers mental states or cognitive states, subjective and objective complexity, flexibility, driver prototypes and traits. For example, if conceptual understanding is important for the application at hand, white box models have a clear advantage, whereas for example when the complexity of the scenarios that need to be addressed grows, black box models, which can be learned from large naturalistic datasets, may be more appropriate.

4.11 Modeling workflows

Antti Oulasvirta (Aalto University, FI), Alexandra Bremers (Cornell Tech – New York, US), and Tuomo Kujala (University of Jyväskylä, FI)

In this focused session, we discussed the need to develop modeling workflows specific for this domain. In contrast to empirical sciences (e.g., [1]), models in human-AV interaction need to be, eventually, deployed in interactive systems or as decision-support systems in design. Applicability gains emphasis. However, despite the practical aims of modeling efforts, it is important to retain a level of theoretical and biological plausibility. It was also observed, that because of the safety-criticality of this domain, simulation studies are preferred prior to empirical studies in the wild.

Based on these and other observations, four design principles for a revised modeling workflow were set out: 1) put iteration to the heart of model development; 2) engage in both simulated and real-world studies, 3) aim for robust, verified modeling outputs, and 4) constant contact with theory (in cognitive sciences, biology, neurosciences).

The attached figure shows the revised modeling workflow proposed by the team.

References
1   Gelman, Andrew, et al. *Bayesian workflow*. arXiv preprint arXiv:2011.01808, 2020.

4.12   What is prediction?

Antti Oulasvirta (Aalto University, FI), Jelmer Borst (University of Groningen, NL), Patrick Ebel (Universität Köln, DE), Martin Fränzle (Universität Oldenburg, DE), and Nele Rußwinkel (TU Berlin, DE)

In this focused session, the topic of how to prioritize possible types of predictions was discussed. Prediction is the essential feature of computational models, and it is essential to move beyond merely isolated tests. (Think: Newell's 1973 paper "You can't play 20 questions with nature and win.").

The group agreed that prediction should not be confused with other uses of models, such as description, explanation, and counterfactual reasoning. These are linked to each other but different. The essence of prediction is that it involves some verifiable statement about a future state or unseen situation, given some observational data. What distinguishes driving as a domain is the need to predict what happens in previously unseen future situations, for example when doing verification of an AI's behavior across different situations, some of which can be very rare.

It was also noted that predictions of driver behavior are needed in different ways in different processes, such as user-centered design, safety engineering, and for AI development. Also, different SAE Levels pose different requirements for prediction; like Level 3 poses a requirement for predicting driver's availability. It was noted that an exercise should be carried out to assess these needs from a cognitive modeling perspective.

It was agreed that there is a multiplicity of predictions needed in this domain. This reflects the fact that computational models are used in widely different driving contexts, with different inputs, purposes, and time-ranges. For example, predicting whether a pedestrian is likely to step on a crosswalk occurs in the timespan of a few seconds, while route prediction occurs in the timespan of minutes to hours. There are micro-level and macro-level predictions that are needed, for example arousal states and states related to social interaction with other traffic users. And these interact. There are predictions of latent states that are not directly measurable but measurable via indirect physiological reaction (skin conductance for example). In real applications, we sometimes need to predict behavioral outcomes in driving, but sometimes we need to estimate a latent factor like the level of workload or stress. The group discussed what would be the most valuable thing to predict. There was some consensus that it should be the future state of the driver+vehicle+scene system within some time horizon. This system includes the driver, the vehicle, and the driving environment including the pedestrians.

The need for academic reseach to consider cross-validation practices was discussed. It was perceived as a harder measure of a model's predictive capability. In statistics-based models, parameters are often fit to the same dataset on which fitness metrics are computed. However, the cross-validation practices that are followed in machine learning may not be directly useful for this domain. For example, we sometimes need to predict what happens with a new driver that has not been included in the training dataset (leave-user-out, or LOU, cross-validation). In many machine learning pipelines, the training dataset might contain samples from each user, thus 'leaking' information in a way that is not plausible from a deployment point-of-view in this domain. It was further remarked that safety-related incidents are rare, which complicates prediction, the acquisition of training/reference data for computational models, as well as validation.

Prediction is almost always done under uncertainty, and uncertainty should be factored in the system's action. Therefore a prediction might be more valuable if it also contains a measure of confidence intervals. The need for a probabilistic representation of outcomes is the greater the further in to the future the prediction tries to reach. However, when it comes to the idea of showing this to the driver, or an uncertainty visualization, the problem is non-trivial. Showing uncertainty to a driver can be confusing, as some earlier studies suggest. On this note, "a prediction is an intervention": showing a prediction to a driver affects the driver's future behavior. In multi-agent systems research, there are some formalisms to describe such interactions, but they are rarely done with computational models of humans.

We also discussed the need for updating the way we do models, i.e .the modeling workflow, with methods like sensitivity analysis and what is called parameter recovery.

4.13 Which models should be used and what should be their content?

Shadan Sadeghian Borojeni (Universität Siegen, DE), Martin Baumann (Universität Ulm, DE), Luisa Heinrich (Universität Ulm, DE), Dietrich Manstetten (Robert Bosch GmbH – Stuttgart, DE), Roderick Murray-Smith (University of Glasgow, GB), and Hatice Sahin (Universität Oldenburg, DE)

In this focused group, we discussed different types of models and their content. Our understanding of a computational model is an executable, precise model that is defined through an unambiguous formulation (mathematical, logical, code, etc.) and is replicable. Computational models can cover different types of models such as cognitive, psychological, social behaviour. To understand what model to use, we first need to define the purpose of the model, e.g., predictive or descriptive. In modelling interaction between humans and automated vehicles, black-box models seem promising. However, the drawback is the lack of knowledge they provide about causal relationships. Therefore, we require to integrate machine learning approaches with causal models, predict the uncertainty of models, and test them in different contexts to understand their validity. This helps us move from black-box data-based models of behaviour towards causal first-principles models, which can generalise better, due to capturing key aspects of the system.

Another topic discussed in this focus group was predicting human behaviour while they adapt to changes in the design of the system. Predicting these emerging behaviours without currently existing data seems challenging. We proposed applying closed-loop models that

allow the integration of the adaptive behaviour of humans. Moreover, descriptive models may help us develop causal models that allow predicting emergent behaviour or future behaviour that we currently do not have a lot of data about.

4.14 Which scenarios should be taken into account in computational models of human-automated vehicle interaction?

Hatice Sahin (Universität Oldenburg, DE), Duncan Brumby (University College London, GB), Jussi Jokinen (University of Jyväskylä, FI), and Shadan Sadeghian Borojeni (Universität Siegen, DE)

We have started by thinking possible scenarios for each level of automation, and soon we realized that we could not possibly list all the scenarios. To have a better structure of the discussion we have focused on categorizing these scenarios from different angles. This brought up whether a taxonomy regarding possible scenarios existed in research. Eventually, we have categorized factors originated by the user and the external world. This way, we presented 2 x 2 table of factors regarding the user and external world in SAE levels 1-4 and in level 5. On SAE levels 1-4 User factors included purpose of commute, driving experience, fatigue and engagement. Trust has been found related to all levels. Level 5 included personal preferences such as sustainability, costs and favorable routes. External or contextual factors were mentioned in all levels. These included social norms, environmental factors such as light and weather conditions and cultural or regional norms and formal rules.

4.15 Modeling Trust in Automation

Philipp Wintersberger (TU Wien, AT), Martin Baumann (Universität Ulm, DE), Justin Edwards (ADAPT Centre – Dublin, IE), Luisa Heinrich (Universität Ulm, DE), and Tuomo Kujala (University of Jyväskylä, FI)

We discussed potential approaches to model the concept of "trust in automation" in the driving domain. In particular, we discussed an experimental design that could allow us to model trust as a function of user interventions and monitoring (i.e., "occlusion" of the driving environment) behavior when a driver cooperates with a lane-keeping system (i.e., an SAE level 2 driving automation system).

Assuming that both monitoring and intervention behavior are related to a users' subjective trust, the experiment could work the following way:
- A participant experiences an imperfect lane keeping system, i.e., the automation keeps the vehicle in the lane center but frequently starts to deviate from the ideal trajectory, slowly towards the lane boundaries.
- We assume that a driver observing this process would at some point start to correct the maneuvering of the vehicle to maintain a trajectory in the lane center again.

- While observing the situation and frequently correcting the steering, the participant needs to complete a (comparably short) trust questionnaire multiple times during the experiment.
- The collected data will then be analyzed to determine if the deviation that a driver allows the vehicle to leave the ideal trajectory is linked to the subjective trust ratings. We hypothesize, that a driver with low trust would correct the vehicles' maneuvers earlier and more frequently, than a driver who trusts the automation more.
- Additionally, we could include a secondary task to extend the principle to monitoring behavior. In other words, a high-trusting driver may monitor the driving automation system less often than a low-trusting driver.
- The data observed in both conditions (i.e., intervening/monitoring) could then be used to model the subjective trust of a driver. Given that the assumptions described above are valid, a drivers' monitoring and intervention behavior can then be used to derive their trust levels.

5   Panel discussions

5.1   Summary of Panel 1: How can models inform design?

Antti Oulasvirta (Aalto University, FI), Alexandra Bremers (Cornell Tech – New York, US), Lewis Chuang (LMU München, DE), Debargha Dey (TU Eindhoven, NL), Andreas Riener (TH Ingolstadt, DE), and Shadan Sadeghian Borojeni (Universität Siegen, DE)

Models are most useful if they are more than abstract, theoretical vehicles. They should not live in a vacuum, but be related to problems and issues in the real world. Therefore, the seminar wanted to discuss how models can inform the design of (in-)vehicle technology, and how they can inform policy. As both of these topics can fill an entire Dagstuhl by themselves, the primary objective was to identify the most pressing issues and opportunities.

Andreas Riener raised the question who we are designing for: the driver, the passenger, or both? The different stakeholders have varying demands for design. He then brought up the topic of cooperation with the OVID model of Roberts (1998).

Shadan Sadeghian followed up with the question on the blurring of boundaries between the vehicle and the human driver, which is further confused by the increasing levels of automation. She then discussed the dfiferent levels and their differing requirements posed to design.

Dave Dey discsused experimental research on automated vehicles and other road users, zooming into his own work on short-sighted aspects of interaction, like how pedestrians respond to a particular event. He raised the issue that research in this domain is complicated by multiple variables and it is hard to design solid research studies that can tease apart causes from confounds.

Alexandra Bremers took an interaction design point-of-view, talkign about how wizard of oz method can be used to study how people respond to automated vehicles, how grounding happens. However, this is confounded by the fact that automated vehicles are conceptually changing: they are not just transportation vehicles but becoming something something more: spaces for coming together, sharing with other people etc.

Antti Oulasvirta discussed design as counterfactual thinking. Consequently, the goal is to understand how design affects the adaptation of human behavior. He talked about emerging theories from ML, especially reinforcement learning, which offers a way to simulate the emergence of adaptive behavior.

Lewis Chuang raised the issue of what a model is, proposing a definition: "A representation of what might be going on "out there" based on our biological perception". This compares to information transer. Such models are important to avoid heuristic reasoning when thinking about automated driving. John Senders pioneering work is a great example of combining theoretical work with modeling and rigorous experimental research.

After the opening statements, the panel and the audience discussed three questions related to this theme:

(1) Types of questions: what types of questions exist at a design and policy level about human-automated vehicle interaction?

(2) How to inform decisions: How can models be used to inform design and policy decisions? What level of detail is needed here? What are examples of good practices?

(3) Integration: Integration can be considered in multiple ways. First, how can ideas from different disciplines be integrated (e.g., behavioral sciences, engineering, economics), even if they have at times opposing views (e.g., monetary gains versus accuracy and rigor)? Second, how can models become better integrated in the design and development process as tools to evaluate prototypes (instead of running empirical tests)? And third, how can models be integrated into the automation (e.g., as a user model) to broaden the automation functionality (e.g., prediction of possible driver actions, time needed to take over)?

As possible future topics, the participants saw three topics rise above others: 1) defining better what models do in design, especially what they predict; 2) finding connections between qualitative and quantitative understandings; and 3) making models more accessible for designers, i.e easier to use and better integrated to their practices.

5.2 Summary of Panel 2: What phenomena and driving scenarios need to be captured in computational models of human-automated vehicle interaction?

Martin Baumann (Universität Ulm, DE), Luisa Heinrich (Universität Ulm, DE), Andrew Kun (University of New Hampshire – Durham, US), Dietrich Manstetten (Robert Bosch GmbH – Stuttgart, DE), Nikolas Martelaro (Carnegie Mellon University – Pittsburgh, US), and Hatice Sahin (Universität Oldenburg, DE)

The second panel discussed what phenomena and driving scenarios need to be captured in computational models of human-automated vehicle interaction. The aim of this panel discussion was, as described by the organizers of the Dagstuhl Seminar, to both advance theory on human-automation interaction while also contributing to understanding realistic case studies for human-automation interaction that are faced for example by industry and governments. The following examples of possible relevant phenomena and driving scenarios were identified during the preparation of the seminar:

1. Transitions of control and dynamic attention: When semi-automated vehicles transfer control of the driving task back to the human, they require accurate estimates of a user's attention level and capability to take control (e.g., [14, 16])
2. Mental models, machine models, mode confusion, and training and skill: Models can be used to estimate human's understanding of the machine and vice-versa (e.g., [15]). Similarly, they might be used to estimate a human driver's skill level, and whether training is desired.
3. Shared control: In all these scenarios, there is some form of shared control. Shared control requires a mutual understanding of human and automation. Computational models can be used to provide such understanding for the automation (e.g., [17]).

These three areas and others were mentioned and discussed during the panel discussion. It started with short pitches by the panelists that are shortly summarized in the following section.

The following attendees were speakers on this panel: Luisa Heinrich, Nikolas Martelaro, Andrew Kun, Dietrich Manstetten, Hatice Sahin, Martin Baumann

Luisa Heinrich emphasized the importance of the take-over scenario as one key area where computational models of driver-automated vehicle interaction could be helpful. The main challenge here is how to design transitions strategies that support a safe and efficient transfer of control from the automation to the human driver, especially in cases when the human driver is out of the loop.

Nikolas Martelaro pointed out the importance of the complexity of traffic scenarios in which human drivers are required to take back control from the vehicle automation. A main target area for computational modelling could be how humans manage their situation awareness in such complex traffic situations given that there are many types of road users that have to be taken into account, such as pedestrians, cyclists, buses, passenger cars. He also emphasized that different kind of drivers need to be considered. Finally he mentioned as a last highly relevant phenomenon that automation might reduce some types of accidents but also leads to new kinds of accidents.

The third pitch was given by **Andrew Kun** and he focussed on the importance of bumper-to-bumper traffic as one relevant research area where computational models might be useful. The other topic Andrew brought into the discussion are the small bursts of interaction that occur during driving with automated vehicles where either the car or the human only contribute to the driving task for a short period of time. Additionally, the different types of tasks that are carried out while driving should be considered as relevant topic, such as non-driving related tasks, visual tasks, etc.

Dietrich Manstetten structured his pitch around for essential questions: What? Where? What for? In which situation? Regarding "what" Dietrich distinguished status vs. behavior models. The question here is what is the input and what is the output of a model. So the behavior could be used as input for status models of driver attention, whereas the current attention allocation could be an input to a computational model of lane change behavior. Regarding the "where" question he raised the points of lab vs. series vehicles. Whereas online restrictions are extremely relevant for series vehicles lab vehicles can work with offline data analysis tasks and have the freedom for highly sophisticated sensors that are nearly impossible in series vehicles. With regard to the "what for" question the main question is what is model used for and the main distinction is between safety and comfort. Which purpose is addressed determines the requirements on the quality of the computational models. Regarding the question "in which situation" Dietrich Manstetten opposed the take-over situation with the automation phase. He made clear that it is highly important to consider

the environment and the situation any model was designed to work in. More aspects of the situation that should be considered are the traffic situation, the subtasks of the driving task, the time horizon or the vehicle type.

Hatice Sahin emphasized the fact that the general traffic ecosystem in which behavior occurs is generally neglected in studies and she raised the question how to integrate this complexity into computational models. Additionally, situational aspects such as traffic density and the presence of especially vulnerable road users need to be integrated in modelling activities.

The last pitch was presented by **Martin Baumann**. He focussed on the question of how humans interact with automated systems in traffic. He pointed out that the problem in addressing this question is in many cases not a lack of data, but a lack of precise definitions of the relevant constructs and predictive models. Computational models might be one promising way therefore to generate significant progress in answering this question. For this he mentioned three points that need to be addressed with computational models: The first is how people construct and maintain situational awareness in dynamic traffic situations as this mental representation is the basis for the decisions humans take. Second, long-term effects of the interaction with automated vehicles are highly important to consider. And third, more complex and realistic situations have to be considered in order to really evaluate the effect of automated systems on human behavior in traffic.

In the subsequent discussion the following topics, among others, where raised:

1. What kind of data can be used as a basis for models and especially how can data from the real world be integrated in the formation of computational models? There are big data sets, especially at companies that would be very helpful in building and validating models if they could be accessed by the scientific community.

2. With regard to the question which phenomena should be addressed with computational models the distinction was made between a technological perspective – what scenarios / environments should be addressed – and a human-centered perspective – what are the most important cognitive processes determining human behavior in traffic across different scenarios. Starting from the situation and the environmental context might allow to identify those cognitive phenomena that are relevant and that should be integrated into computational models of human behavior in these situations. On the other hand the point was made that it might be quite difficult to identify those scenarios that are really relevant in human-automated vehicle interaction as automated vehicles are still developing and the data we currently collect and use might not be relevant ones in the future. Additionally, starting from specific situations might lead to a situation where different theories and models for different situations are developed that are not compatible and cannot be generalized across different contexts. With regard to this the need for a unified and integrated modeling approach was formulated.

3. As one important situational factor the complexity of the situation was mentioned, and here it was stressed that it is not so much the "objective" complexity that is relevant, but the subjectively experienced complexity.

4. The topic of trust into technology and how it could be modelled computationally was discussed in depth. The different types of models and theories that the participants of the discussion use to investigate and measure trust were collected during the discussion. Relevant phenomena could be overtrust and distrust, uncertainty as a possible mechanism underlying trust.

5. One of the phenomena that was also mentioned in the context of the trust discussion was learning as a highly relevant phenomenon for computational models of human-automated

vehicle interaction. The experience with automated vehicles will shape the future behavior and currently many models of human-automated vehicle interaction are based on first time encounters with the technology.

The following papers where mentioned during the discussion and listed in the notes of the discussion:

References

1   Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An Integrative Model of Organizational Trust. The Academy of Management Review, 20(3), 709–734. https://doi.org/10.2307/258792

2   Teodorovicz, T., Kun, A. L., Sadun, R., & Shaer, O. (2022). Multitasking while driving: A time use study of commuting knowledge workers to assess current and future uses. International Journal of Human-Computer Studies, 162, 102789. https://doi.org/10.1016/j.ijhcs.2022.102789

3   Markkula, G., & Dogar, M. (2022). How accurate models of human behavior are needed for human-robot interaction? For automated driving?. arXiv preprint arXiv:2202.06123.

4   Sibi, S., Balters, S., Fu, E., Strack, E., Steinert, M, & Ju, Wendy. (2020). Back to School: Impact of Training on Driver Behavior and State in Autonomous Vehicles. 10.1109/IV47402.2020.9304537.

5   Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y. C., de Visser, E. J., & Parasuraman, R. (2011). A Meta-Analysis of Factors Affecting Trust in Human-Robot Interaction. Human Factors, 53(5), 517–527. https://doi.org/10.1177/0018720811417254

6   Khavas, Z. R. (2021). A Review on Trust in Human-Robot Interaction. arXiv preprint arXiv:2105.10045.

7   Khavas, Z. R., Ahmadzadeh, S. R., & Robinette, P. (2020, November). Modeling trust in human-robot interaction: A survey. In International Conference on Social Robotics (pp. 529-541). Springer, Cham.

8   Hoff, K. A., & Bashir, M. (2015). Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust. Human Factors, 57(3), 407–434. https://doi.org/10.1177/0018720814547570

9   Johnson, D., & Grayson, K. (2005). Cognitive and affective trust in service relationships. Journal of Business Research, 58(4), 500–507. https://doi.org/10.1016/S0148-2963(03)00140-1

10 Lee, J. G., & Lee, K. M. (2022). Polite speech strategies and their impact on drivers' trust in autonomous vehicles. Computers in Human Behavior, 127, 107015.

11 Jussi P. P. Jokinen and Tuomo Kujala. 2021. Modelling Drivers' Adaptation to Assistance Systems. In 13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '21). Association for Computing Machinery, New York, NY, USA, 12–19. https://doi.org/10.1145/3409118.3475150

12 Kujala, T., & Lappi, O. (2021). Inattention and Uncertainty in the Predictive Brain. Frontiers in Neuroergonomics, 2, 718699. https://doi.org/10.3389/fnrgo.2021.718699

13 Azevedo-Sa, H., Jayaraman, S. K., Esterwood, C. T., Yang, X. J., Robert, L. P., & Tilbury, D. M. (2021). Real-Time Estimation of Drivers' Trust in Automated Driving Systems. International Journal of Social Robotics, 13(8), 1911–1927. https://doi.org/10.1007/s12369-020-00694-1

14 Janssen, C. P., Iqbal, S. T., Kun, A. L., & Donker, S. F. (2019). Interrupted by my car? Implications of interruption and interleaving research for automated vehicles. International Journal of Human-Computer Studies, 130, 221-233.

15 Janssen, C. P., Boyle, L. N., Kun, A. L., Ju, W., & Chuang, L. L. (2019). A hidden markov framework to capture human–machine interaction in automated vehicles. International Journal of Human–Computer Interaction, 35(11), 947-955.

16 Wintersberger, P., Schartmüller, C., & Riener, A. (2019). Attentive User Interfaces to Improve Multitasking and Take-Over Performance in Automated Driving: The Auto-Net of Things. International Journal of Mobile Human Computer Interaction, 11(3), 40-58.

17 Yan, F., Eilers, M., Weber, L., & Baumann, M. (2019). Investigating Initial Driver Intention on Overtaking on Rural Roads. In 2019 IEEE Intelligent Transportation Systems Conference (ITSC) (pp. 4354-4359).

5.3   Summary of Panel 3: What technical capabilities do computational models need to possess?

Antti Oulasvirta (Aalto University, FI), Jelmer Borst (University of Groningen, NL), Martin Fränzle (Universität Oldenburg, DE), Myounghoon Jeon (Virginia Polytechnic Institute – Blacksburg, US), Otto Lappi (University of Helsinki, FI), Gustav Markkula (University of Leeds, GB), and Nele Rußwinkel (TU Berlin, DE)

The panel discussed technical capabilities required of computational models in this domain. Although the nature of different modeling frameworks and different studies might differ, what do we consider the core functionality? For example, there are capabilities related to:

(1) Compatibility: To what degree do models need to be compatible with simulator software (e.g., to test a "virtual participant"), hardware (e.g., be able to drive a car on a test track), and other models of human thinking?

(2) Adaptive nature: Computational models aim to strike a balance between precise predictions for more static environments and being able to handle open-ended dynamic environments (like everyday traffic). How can precision be guaranteed in static and dynamic environments? How can models adapt to changing circumstances?

(3) Speed of development and broader adoption: The development of computational models requires expertise and time. How can development speed be improved? How can communities benefit from each other's expertise?

The panel consisted of the following speakers: Gustav Markkula, Martin Fraenzle, Myounghoon Jeon (Philart), Jelmer Borst, Otto Lappi, and Nele Russwinkel.

Gustav Markkula discussed the compatibility of software and hardware, which is critical at the very end of an applied project, but much less considered earlier in the model developemnt stage. It is therefore critical to study the ultimate application context at least roughly. Similarly from an applied perspective, the speed of development and the flexibility of modelling is important. Models should be extensible with new data, new features, new capabilities etcetera.

Martin Fraenzle followed up and discussed how the idea of running models "in situ" might be out of scope at the moment.

Myounghoon Jeon brought up Distract-R as a great example of modeling, which allows different parameters to be simulated and rapid prototyping while being hardware and software agnostic and yet have some ecological validity. He echoed the need for adaptability and extensibiliy. He further brought up the topic of motions, which is presently missing from cognitive models although clearly important.

Jelmer Borst discussed the desiredata of modeling, one of which is that they should be able to drive the car. However, ACT-R is far from that goal according to the panelist. It is too brittle. There is a need for hybrid models that combine cognitive architectures and machine learning to overcome this issue. This may allow a cognitive model to operate in the background while the user is driving and warn users if something out of line is taking place, such as overloading.

Otto Lappi proposed a more philosophical perspective. He claimed that the first question to ask is who you are modeling for: the designer, the engineer, the scientist, or someone else? This has significant implications to the capabilities we want from models, and also our validation criteria.

Nele Russwinkel followed up on the topic of what is required of a model during driving, pointing out that ACT-R may be useful if it can anticipate the driver, even if now able to drive the car. She pointed out that latent variables, such as those related to situation awareness, are important for a car to understand. However, presentday models mostly do not afford this in a real-time system, due to computational intensity. She pointed out that a lot of modeling is still missing when it comes to the different levels of automation in vehicles. Moreover, some of the events we are interested in are rare and therefore inherently difficult to predict.

Discussion with the audience followed up on the need for better simulation environments and hybrid models that use ML. There was disagreement on whether simulation environments are too simplistic and can or cannot contain the richness of perceptual and other cues that drivers exploit in driving. The audience gravitated toward the need for developing a roadmap for cognitive models. Participants agreed that one of the outstanding goals is to understand the different SAE levels better, especially from the perspective of their psychological consequences and, therefore, their consequences on cognitive models.

Future topics that were raised included:

1) How to allow drivers update models about themselves?

2) How to trade off predictive accuracy and computational costs in such a way that we can allow cognitive models to run in real-time systems?

3) How to go from mental models research to scene understanding, which is critical in driving.

5.4 Summary of Panel 4: How can models benefit from advances in AI while avoiding its pitfalls?

Christian P. Janssen (Utrecht University, NL), Duncan Brumby (University College London, GB), Birsen Donmez (University of Toronto, CA), Justin Edwards (ADAPT Centre – Dublin, IE), Mark Eilers (Humatects – Oldenburg, DE), Moritz Held (Universität Oldenburg, DE), Jussi Jokinen (University of Jyväskylä, FI), and Roderick Murray-Smith (University of Glasgow, GB)

The fourth panel discussed how computational models that are developed for Human-Automated Vehicle Interaction can benefit from advances in AI while also avoiding some of its pitfalls. The summary below is made by Chris Janssen, building on crowd-sourced community notes of the sessions and snippets of panel members' presentations or text where possible.

Before the meeting, the organizers had identified that there are many developments in AI that computational models can benefit from. Three examples are advances in (1) simulator-based inference (e.g., [1]) to reason about possible future worlds (e.g., varieties of traffic environments), (2) reinforcement learning [2] and its application to robotics [3] and human driving [4], and (3) deep learning [5] and its potential to predict driver state or behavior from sensor data. At the same time, incorporation of AI techniques also comes with challenges that need to be addressed. Three potential challenges are for example:

1. Explainability: Machine learning techniques are good at classifying data, but do not always provide insight into why classifications are made. This limits their explainability and is at odds with the objective of computational models to gain insight into human behavior. How can algorithms' explainability be improved?
2. Scalability and generalization: How can models be made that are scalable to other domains and that are not overtrained on specific instances? How can they account for future scenarios where human behavior might be hard to predict [6]?
3. System training and corrective feedback: if models are trained on a dataset, what is the right level of feedback to correct an incorrect action to the model? How can important new instances and examples be given more weight to update the model's understanding without biasing the impact?

During the panel, these and other themes were discussed. We started with short pitches by the panelists, which are summarized first.

Roderick Murray-Smith discussed among others the following points.

(1) A priori insight can be placed into machine learning models in multiple ways. Forward and inverse modelling components were discussed as being particularly relevant.

(2) The advent of deep learning models has helped to place more complex models of perception into control models.

(3) This has knock-on effects of explainability: is trying to explain a complex policy model a good idea? Or is it better to ensure that the model learns a value function that is relevant to the human and then optimized?

(4) Engineering approaches often follow a modular approach, where multiple components are solved from first-principles, before being placed in bigger systems. The question is whether such an approach is scalable to a more complex context such as driving, where a closed -loop simulation might not be feasible.

Mark Eilers discussed that the best models for driving an automated vehicle do not necessarily need to be models that drive in a human way, or based on a human-like model. Mark saw potentials and possibilities in the following fields and areas.

(1) robotics, for example imitation learning,

(2) reinforcement learning and optimal control,

(3) deep learning, for example the use of new unsupervised learning techniques.

All these techniques can be adapted to fit existing models and contexts. Mark identified as pitfalls that there might be an identifiability problem: data alone can't always tell the researcher which of the many different approaches or models is the correct one. Theory and experiments together must inform the model and the modeling community.

Duncan Brumby reflected on his own past work in preparation for the meeting. When he modeled driving 15 years ago, he used a relatively simple model, but it was useful and valuable (e.g., [7]). At the time, there was talk of how reinforcement learning models can one day give even more insight, and it is good to see that that potential has come to fruition (e.g., [4]). Duncan now sees three areas where science of AI has lead to great developmens and still has potential to grow further.

(1) Models of the driver: science can help to create better models
(2) Models for the car itself: current semi-automated vehicles have lots of technology and models in it, for example to go around a corner and break when a car in front of it stops. However, the models for these tasks are developed by companies, and the question is how it will reach science.
(3) The technology with which people interact in the car (e.g., secondary devices) have evolved. The tools that people use are different.

Overall, due to these changes, some of the parameters that underly older models might need updating. For example, when do we call something distracting or not differs between settings.

Birsen Donmez looked at the topic of this panel from a statistics and experimentation perspective. Some machine learning approaches such as Deep Learning are very powerful, as they enable to model complex non-linear systems. Capturing multiple predictors and their interactions go beyond the model typical traditional statistical or cognitive models. However, they do require big data sets or samples. More traditional statistical tools do not always require such larger sets, but can also work on smaller samples. There are challenges for both the big data set (and machine learning) and small data set (and statistics) approaches. For the bigger data sets the general challenge is that these are often collected by industry and not available for researchers. How can this be changed? When we collect our own datasets they might also not always be big enough, and when using data sets of others, that might also not always be trivial. For smaller data sets there is a limited sample problem, when applying machine learning techniques. It requires insight to for example split the dataset appropriately into training set and test set. Birsen sees potential for situations where researchers do get hold of bigger data sets. They have the right background to be critical about the data. For example, to ask questions such as: is the data representative? How was data collected? Do correlations exist?

Jussi Jokinen discussed a model of how humans can interact with an automated vehicle. A cooperative AI can for example observe human user behavior, and then try to identify posterior probabilities of user states (in a Bayesian fashion). There are however potential pitfalls in that such a Bayesian inference process can be slow and the results of a modeling framework are sometimes surprising and don't seem plausible. Given that the model is not always correct, it is useful to explicitly consider (un-)certainty. In Jussi's own work (e.g., [4]), he benefits from applying the computational rationality framework (e.g., [8, 9]). Such a model expresses how latent variables impact behavior. A reinforcement learning agent is developed that works within constraints of a cognitive architecture and task environment to come up with the optimal policy that if architecture and task are correctly defined should be similar to user's behavior. It is cool when it works, but there are lots of moving parts. Jussi sees opportunities (and challenges) in refinement of the ability to use inference techniques in these models: how can one be confident that the learned model is indeed the correct model?

Justin Edwards drew parallels with the history of language models and his experience in the Conversational User Interfaces (CUI) community. Coming up with language generation and processing models long proved a challenge to the field. Initially there were many rule-based models. The last decade has seen a wider range of models that use large datasets and are data-driven. Data-driven models have surpassed the more theory driven models in quite a lot of tasks. However, they also are occasionally deeply flawed. The samples on which training takes place are not always general, but are used for generalization purposes. This showed for example in the Delphi system, which was meant to be capable of moral judgements, but used data from reddit chatboards for training.So, the big lesson learned for the automotive community is: as new technology becomes available, be aware of where the data comes from. What are the biases? Those might carry through in surprising ways.

Moritz Held focused mostly on two themes. First: explainability. Hybrid models that combine machine learning with theory-driven models might be fruitful in this regard. For example, combining ACT-R models with techniques from Bayesian spatial networks. The second topic was scalability of models. Moritz was surprised by some discussions in previous days regarding scalability. He would expect that there is always some robustness in a model. There is no need for (for example) and ACT-R supermodel to handle every driving situation. But, he would expect it to at least handle different driving scenarios.

Chris Janssen talked about four points. First, it is important to think about function allocation: how are tasks divided between humans and the automated vehicles. Models can play a role in this regard (see e.g. [10]). Second, in such discussions, it is tempting to think of simple heuristics such as "men are better at [some tasks] and machines are better at [other tasks]" [11]. However, humans and machines are dynamic (they can both learn and adapt), and they operate in varying contexts. Therefore, static heuristics might be useful at first, but might not always be appropriate for these dynamic contexts. Cognitive models should take such dynamics into account. Third, explainability is becoming increasingly important inside and outside of science. Computational models can tie in with this by combining theory-driven and data-driven models. In other words: it is important that models are connect to theory (not just data), but also that theory is connected to practice.

During the discussion that followed, we talked among others about the following aspects:
1. What creates distributional shifts in models? How can models be made more robust against this?
2. How can the designers and researchers of modelers better communicate what the goals and limitations of models are? For example, for which level of automation they are designed.
3. There seems to be a balance between generalization of a model versus having a more structured / fixed model. Models differ in which aspects they keep fixed (as assumptions, framework, or architecture) and which aspects they train or learn (for example, based on data).
4. There was also a discussion about how computational models can be designed such that designers and developers (of cars) can better benefit from them. There is a potential discrepancy here in that designers want to test multiple, flexible designs whereas models are often trained on or designed for more fixed tasks.
5. There is a need to be more systematic about the model selection process.
6. There is an open question about for what types of research questions small or large datasets are needed.

References
1   Kangasrääsiö, A., Jokinen, J. P., Oulasvirta, A., Howes, A., & Kaski, S. (2019). Parameter inference for computational cognitive models with Approximate Bayesian Computation. *Cognitive Science*, 43(6)
2   Sutton, R., & Barto, A. G. (2018). *Reinforcement learning: An introduction.* Cambridge, MA: MIT Press
3   Levine, S. (2018). *Reinforcement learning and control as probabilistic inference: Tutorial and review.* arXiv preprint arXiv:1805.00909.
4   Jokinen, J. P., Kujala, T., & Oulasvirta, A. (2021). Multitasking in driving as optimal adaptation under uncertainty. *Human factors*, 63(8), 1324-1341.
5   Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning.* Cambridge, MA: MIT press.

6 Bainbridge, L. (1983). Ironies of automation. In *Analysis, design and evaluation of man–machine systems* (pp. 129-135). Pergamon.

7 Brumby, D. P., Salvucci, D. D., & Howes, A. (2009, April). Focus on driving: How cognitive constraints shape the adaptation of strategy when dialing while driving. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1629-1638).

8 Lewis, R. L., Howes, A., & Singh, S. (2014). Computational rationality: Linking mechanism and behavior through bounded utility maximization. *Topics in cognitive science*, 6(2), 279-311.

9 Oulasvirta, A., Kristensson, P. O., Bi, X., & Howes, A. (Eds.). (2018). *Computational interaction*. Oxford University Press.

10 Janssen, C. P., Boyle, L. N., Kun, A. L., Ju, W., & Chuang, L. L. (2019). A hidden markov framework to capture human–machine interaction in automated vehicles. *International Journal of Human–Computer Interaction*, 35(11), 947-955.

11 Fitts PM (ed) (1951) *Human engineering for an effective air navigation and traffic control system*. National Research Council, Washington, DC

5.5 Summary of Panel 5: What insights are needed for or from empirical research?

Shamsi Tamara Iqbal (Microsoft – Redmond, US), Linda Ng Boyle (University of Washington – Seattle, US), Benjamin Cowan (University College – Dublin, IE), Patrick Ebel (Universität Köln, DE), Wendy Ju (Cornell Tech – New York, US), Tuomo Kujala (University of Jyväskylä, FI), Philipp Wintersberger (TU Wien, AT), and Fei Yan (Universität Ulm, DE)

Computation models are only as good to the degree that they can describe and predict phenomena in the real world. In particular, many current computational models capture the results of a single experiment based on the current context of what is being modeled. However, behavior might change with more exposure to and experience with automated technology. To make models most useful over time, empirical testing is of paramount importance, especially in order to evaluate models that look at current behavior, as well as models that are based on projections of future behavior. The panel on empirical research insights focused on the question above, as well as related questions such as how to study phenomena where computational models are not commercially available, and how to map findings from simulator tests to real world scenarios.

Patrick Ebel started the conversation with a focus on mapping real world data to input for computational models. Clearly there is a need for models to be used in industry and leveraging existing models can save time and money but there seems to be a barrier in adaptation. A couple of possible causes could be accuracy of models, especially since real world application needs to consider many confounding factors which cannot be typically captured in experimental models. Patrick also raised the question of the comparability of data that come from different simulators and how to reconcile for those differences.

Wendy Ju discussed the generalizability of empirical studies that are done in controlled settings in the lab – where the studies themselves are challenging to run and have additional limitations (e.g. cultural, temporal variation). She proposed the alternative of gleaning results from understanding naturalistic data from real-world driving. However, this will

require instrumentation of vehicles where there needs to be a common agreement of what data is collected, when and how. The data itself can be a contribution to the field, and once we have these datasets, ML can be used to scale up and detect patterns in the diverse data sets that are collected.

Fei Yan talked about how to extract insights from empirical research and black box models and posed the core question of determining when a model is really needed to solve a research question. Empirical research can help understand causal relationships, and black box models can look at theory (e.g. social psychology on human-human interaction) which can be then used to form hypotheses that can be tested in empirical research. Fei proposed using empirical research to understand relevant factors before making the black box model. An example scenario could be the use of a predictive model in a lane change support system that can adapt to a driver's states of uncertainty. But it is difficult to know what models are most applicable in these scenarios. For some domains, ML or Bayesian models might be enough, but depending on the level of accuracy needed, those might not be most appropriate for other domains. Fei also talked about the small sample size in empirical research which helps with deeper dives into relationships between factors.

Tuomo Kujala shifted gears towards modelling driver attention. Driver attention monitoring is required for cars in level 3 or even up to level 4 in order for a car to respond safely and effectively to takeover requests. The challenge is that most attention systems monitor the inside of the car and how the driver's attention is impacted by the internal environment that can be detected. However, a driver's attention can be impacted by factors that are not detectable directly – for example, how much 'free attention' or cognitive capacity is available for the driver and whether it is enough to handle an upcoming scenario in driving. Appropriate attention modelling can decrease the uncertainty. Tuomo proposed that we need prescriptive or normative computational models, not only descriptive models.

Phillipp Wintersberger focused on the specific scenario of what is needed to make Level 3 automated driving possible and presented the argument that simulator results may not be as limited as we assume them to be. The trade off for a data-driven model is that a huge amount of naturalistic data is needed, connecting back to Wendy's point – we as a field do not have established standards on how to get this data. For example, doing a study in a test track may not yield useful data if the test track is not similar to the real world, but similar studies can be designed to be more similar to real world scenarios. Phillipp also talked about how we can get massive data to investigate development of trust in such systems. Is gamification a possible path? He also questioned the need for really big data sets and whether smaller datasets can be combined to get bigger sets.

Shamsi Iqbal talked about the target user – who we are designing for, the current user or the future user who will have different technological capabilities, exposure and experience. Can we effectively extrapolate from the current user to project what the future user will look like? Empirical research typically falls into three buckets – lab experiments, model building approach and naturalistic driving. While naturalistic driving gives the opportunity to learn more about what challenges people face in real world scenarios, there are also privacy and security concerns around collecting massive amounts of data from people. For lab experiments in the other hand, we have the opportunity to test innovative ideas that push the boundary of what might be possible. The challenge is then to determine how to generalize the findings to real world scenarios, where we may not even know what the future scenarios might be.

Linda Boyle brought up the counter point that model-based and empirical research are not mutually exclusive and naturalistic and simulated data should not be mutually exclusive either. Research needs data from separate dimensions, not just comparative studies. Even in

a single lab – different types of data can be collected. The goal should not be just to create a data repository, but also to use results for scenarios such as Bayesian inferences/priors for future models and experiments. Linda emphasized that while field studies are great there is a lot of variation. Lab studies are a must – especially to test rate situations and edge cases that we may rarely or never see in real world data, but still need to be accounted for in our models.

Ben Cowan emphasized that the key to empirical research is observation of phenomena. In his view, the role of empirical research in automated vehicles are: a) to assess how concepts we deem important to automated driving (e.g. situational awareness, attention) manifest in and are impacted within the experience of automated driving. We can learn a lot from conducting qualitative studies on automated driver experience (e.g. observation studies, interviews) to assess the importance of concepts like trust, potential challenges and barriers to adoption etc. Quantitative empirical work, particular lab studies can also be used to test how specific concepts are causally influenced by design or events during the drive to get a clearer picture of what are the key areas for us to focus on (e.g. the design of communication strategies for pedestrians- Lanzer et al, AutoUI 2020 on politeness strategies positive effect on trust and acceptance). They also allow us to experiment with identifying cognitive concepts of importance.

To support the notion of what to model and to give our initial models a test run: Models should be related to the real world. But we need to perhaps incubate these first. Empirical lab based work is important to determine what is and is not important in that real world for the concepts we deem important to research. That is not to say that we cannot and should not do lab-based work, devise, test and assess models for concepts on lab based data to determine how the cognitive concepts may behave within more real world driving contexts.

A key consideration is also in the measurement quality for measuring the concepts we wish to measure. In particular when using questionnaires to assess concepts like trust, it is important that we stay true to any base questionnaires used or conduct the design of our own psychometrics to ensure what we are using are valid, reliable and sensitive to the concepts we measure. On a wider point as a community of researchers we must also ensure that the empirical work we are doing is replicable and open. We need to create an infrastructure and reward system that allows us to embed replication in our empirical research activities (either through replication when publishing the initial study, or in specific tracks at our key conferences and journal publication venues).

In the discussion, Patrick talked about triangulating empirical research and model based research using the following approach: 1) Use model based systems that are built upon naturalistic data. Can help to see patterns in the data and suggest empirical tests, and 2) Use models as a sandbox and test bed approach. How good does a model need to be? If model already shows a study is not worth the time, might be useful to test. Another important topic that emerged from the panel discussion is whether a 'replicate' track should be added to the Auto-UI conference. Key questions that arose included what is the definition of replication, what is the expected rigor, how to address the differences in the experimental setup that might impact the replication, how to share data, how to present the results at the conference and what kind of benchmarks should be setup.

As possible future topics, the following points were extracted from the panel discussion:
- How to connect models even better to needs of industry? They want them, but don't always use them. Why? (e.g., Patrick)
- How to move towards "data is a contribution" and replication (Wendy)
- Do we need models even? (e.g., Fei)

- Prescriptive models are needed; not just descriptive (Tuomo)
- To understand attention, more is needed than just in-car glances (Tuomo)
- Do we need realistic studies? Or is simulator sometimes better (Philipp)
- Opportunity of gamification (Philipp)
- Need for a place to store online data (Philipp)
- Who are we designing tech / experiment / model for? Today's user or tomorrow's user? (Shamsi)
- How to make tools accessible to wider community? For models and tools and experiments. (Linda)
- How to motivate and reward replication systems? (Ben)

6   Open problems

6.1   Relevant papers for modeling human-automated vehicle interaction

Christian P. Janssen (Utrecht University, NL)

License Creative Commons BY 4.0 International license
© Christian P. Janssen

Among the attendees we also gathered an overview of papers that they thought were interesting for researchers in the field.

Before the conference, we sent out a google form and asked attendees to submit papers that they thought were interesting and either written by themselves or others. The suggested papers are for example relevant domains for modeling, or examples of modeling papers. The following papers were suggested: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46]

During the conference, we also crowdsourced a collection of relevant papers. The large majority of these papers contain examples of models or (conceptual) frameworks or datasets that are used for or inspired by models. These papers were suggested: [3, 13, 15, 16, 37, 45, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76]

References

1 Anderson, J. R., Zhang, Q., Borst, J. P., & Walsh, M. M. (2016). The discovery of processing stages: Extension of Sternberg's method. Psychological review, 123(5), 481.

2 Bianchi Piccinini, G., Lehtonen, E., Forcolin, F., Engström, J., Albers, D., Markkula, G., Lodin, J., & Sandin, J. (2020). How do drivers respond to silent automation failures? Driving simulator study and comparison of computational driver braking models. Human factors, 62(7), 1212-1229.

3 Blum, S., Klaproth, O., & Russwinkel, N. (2022). Cognitive Modeling of Anticipation: Unsupervised Learning and Symbolic Modeling of Pilots' Mental Representations. Topics in Cognitive Science, 10.1111/tops.12594. Advance online publication. https://doi.org/10.1111/tops.12594

4 Cummings, M. L., Li, S., & Zhu, H. (2022). Modeling operator self-assessment in human-autonomy teaming settings. International Journal of Human-Computer Studies, 157, 102729.

5 Damm, W., Fränzle, M., Lüdtke, A., Rieger, J. W., Trende, A., & Unni, A. (2019, June). Integrating neurophysiological sensors and driver models for safe and performant automated vehicle control in mixed traffic. In 2019 IEEE Intelligent Vehicles Symposium (IV) (pp. 82-89). IEEE.

6 Eilers, M., Fathiazar, E., Suck, S., Twumasi, D. (2019) Dynamic Bayesian networks for driver-intention recognition based on the traffic situation. In Cooperative Intelligent Transport Systems: Towards high-level automated driving, , pp. 465-495

7 Funk Drechsler, M., Peintner, J. B., Seifert, G., Huber, W., & Riener, A. (2021, September). Mixed Reality Environment for Testing Automated Vehicle and Pedestrian Interaction. In 13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (pp. 229-232).

8 van der Heiden, R. M., Kenemans, J. L., Donker, S. F., & Janssen, C. P. (2021). The effect of cognitive load on auditory susceptibility during automated driving. Human factors, 0018720821998850.

9 Janssen, C. P., Boyle, L. N., Ju, W., Riener, A., & Alvarez, I. (2020). Agents, environments, scenarios: A framework for examining models and simulations of human-vehicle interaction. Transportation research interdisciplinary perspectives, 8, 100214.

10 Janssen, C. P., Iqbal, S. T., Kun, A. L., & Donker, S. F. (2019). Interrupted by my car? Implications of interruption and interleaving research for automated vehicles. International Journal of Human-Computer Studies, 130, 221-233.

11 Jeon, M., Zhang, Y., Jeong, H., P. Janssen, C., & Bao, S. (2021, September). Computational Modeling of Driving Behaviors: Challenges and Approaches. In 13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (pp. 160-163).

12 Jokinen, J. P., & Kujala, T. (2021). Modelling Drivers' Adaptation to Assistance Systems. In 13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (pp. 12-19).

13 Jokinen, J. P., Kujala, T., & Oulasvirta, A. (2021). Multitasking in driving as optimal adaptation under uncertainty. Human factors, 63(8), 1324-1341.

14 Kambhampati, S. (2020). Challenges of human-aware AI systems. AI Magazine, 41(3), 3-17. https://doi.org/10.1609/AIMAG.V41I3.5257

15 Kanaan, D., Ayas, S., Donmez, B., Risteska, M., & Chakraborty, J. (2019). Using naturalistic vehicle-based data to predict distraction and environmental demand. International Journal of Mobile Human Computer Interaction (IJMHCI), 11(3), 59-70.

16 Klaproth, O.W., Halbrügge, M., Krol, L.R., Vernaleken, C., Zander, T.O., & Russwinkel, N. (2020), A Neuroadaptive Cognitive Model for Dealing with Uncertainty in Tracing Pilots' Cognitive State. Topics in Cognitive Science, 12(3), 1012-1029. doi:10.1111/tops.12515

17 Ko, S., Kutchek, K., Zhang, Y., & Jeon, M. (2022). Effects of non-speech auditory cues on control transition behaviors in semi-automated vehicles: Empirical study, modeling, and validation. International Journal of Human–Computer Interaction, 38(2), 185-200.

18 Krishnan, R., & Tickoo, O. (2020). Improving model calibration with accuracy versus uncertainty optimization. Advances in Neural Information Processing Systems, 33, 18237-18248.

19 Kuen, J., Schartmüller, C., & Wintersberger, P. (2021, September). The TOR Agent: Optimizing Driver Take-Over with Reinforcement Learning. In 13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (pp. 47-52).

20 Lavin, A., Zenil, H., Paige, B., Krakauer, D., Gottschlich, J., Mattson, T., ... & Pfeffer, A. (2021). Simulation Intelligence: Towards a New Generation of Scientific Methods. arXiv preprint arXiv:2112.03235.

21 Lechner, M., Hasani, R., Amini, A., Henzinger, T. A., Rus, D., & Grosu, R. (2020). Neural circuit policies enabling auditable autonomy. Nature Machine Intelligence, 2(10), 642-652.

22 Lee, J. Y., & Lee, J. D. (2019). Modeling microstructure of drivers; task switching behavior. International Journal of Human-Computer Studies, 125, 104-117.

23 Leon, J. F. (2020, July). Robust Scene Understanding via Real-Time Approximate Bayesian Computations. In 2020 International Conference on Systems, Signals and Image Processing (IWSSIP) (pp. 13-13). IEEE.

24 Manstetten, D., Beruscha, F., Bieg, H. J., Kobiela, F., Korthauer, A., Krautter, W., & Marberger, C. (2020, October). The Evolution of Driver Monitoring Systems: A Shortened Story on Past, Current and Future Approaches How Cars Acquire Knowledge About the Driver's State. In 22nd International Conference on Human-Computer Interaction with Mobile Devices and Services (pp. 1-6).

25 Markkula, G., & Dogar, M. (2022). How accurate models of human behavior are needed for human-robot interaction? For automated driving?. arXiv preprint arXiv:2202.06123.

26 Martelaro, N., Teevan, J., & Iqbal, S. T. (2019, May). An exploration of speech-based productivity support in the car. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (pp. 1-12).

27 Martín, J. A. Á., Gollee, H., Müller, J., & Murray-Smith, R. (2021). Intermittent control as a model of mouse movements. ACM Transactions on Computer-Human Interaction (TOCHI), 28(5), 1-46.

28 Morando, A., Gershon, P., Mehler, B., & Reimer, B. (2021). A model for naturalistic glance behavior around Tesla Autopilot disengagements. Accident Analysis & Prevention, 161, 106348.

29 Oulasvirta, A. (2019). It's time to rediscover HCI models. Interactions, 26(4), 52-56.

30 Oulasvirta, A., Kristensson, P. O., Bi, X., & Howes, A. (Eds.). (2018). Computational interaction. Oxford University Press.

31 Pakdamanian, E., Sheng, S., Baee, S., Heo, S., Kraus, S., & Feng, L. (2021, May). Deeptake: Prediction of driver takeover behavior using multimodal data. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (pp. 1-14).

32 Pakdamanian, E., Sheng, S., Baee, S., Heo, S., Kraus, S., & Feng, L. (2021, May). Deeptake: Prediction of driver takeover behavior using multimodal data. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (pp. 1-14).

33 Petermeijer, S. M., Tinga, A., Jansen, R., de Reus, A., & van Waterschoot, B. (2021, September). What Makes a Good Team?-Towards the Assessment of Driver-Vehicle Co-operation. In 13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (pp. 99-108).

34 Rehman, U., Cao, S., & MacGregor, C. (2019). Using an integrated cognitive architecture to model the effect of environmental complexity on drivers' situation awareness. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 63, No. 1, pp. 812-816). Sage CA: Los Angeles, CA: SAGE Publications.

35 Ringfort-Felner, R., Laschke, M., Sadeghian, S., & Hassenzahl, M. (2022). Kiro: A Design Fiction to Explore Social Conversation with Voice Assistants. Proceedings of the ACM on Human-Computer Interaction, 6(GROUP), 1-21.

36 Sadeghian, S., Hassenzahl, M., & Eckoldt, K. (2020, September). An exploration of prosocial aspects of communication cues between automated vehicles and pedestrians. In 12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (pp. 205-211).

37 Scharfe, M. & Russwinkel, N. (2019). Towards a Cognitive Model of the Takeover in Highly Automated Driving for the Improvement of Human Machine Interaction. In Proceedings of the 17th International Conference on Cognitive Modelling, Montreal, Canada (pp. 210-215).

38 Suo, S., Regalado, S., Casas, S., & Urtasun, R. (2021). Trafficsim: Learning to simulate realistic multi-agent behaviors. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 10400-10409).

39 Teodorovicz, T., Kun, A. L., Sadun, R., & Shaer, O. (2022). Multitasking while driving: A time use study of commuting knowledge workers to assess current and future uses. International Journal of Human-Computer Studies, 162, 102789.

40 Todi, K., Bailly, G., Leiva, L., & Oulasvirta, A. (2021, May). Adapting user interfaces with model-based reinforcement learning. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (pp. 1-13).

41 Tonolini, F., Radford, J., Turpin, A., Faccio, D., & Murray-Smith, R. (2020). Variational inference for computational imaging inverse problems. Journal of Machine Learning Research, 21(179), 1-46.

42 Wintersberger, P., Schartmüller, C., Shadeghian-Borojeni, S., Frison, A. K., & Riener, A. (2021). Evaluation of imminent take-over requests with real automation on a test track. Human factors, 00187208211051435.

43 Wong, P. N., Brumby, D. P., Babu, H. V. R., & Kobayashi, K. (2019). Voices in Self-Driving Cars Should be Assertive to More Quickly Grab a Distracted Driver's Attention. In Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (pp. 165-176).

44 Wu, T., Martelaro, N., Stent, S., Ortiz, J., & Ju, W. (2021). Learning When Agents Can Talk to Drivers Using the INAGT Dataset and Multisensor Fusion. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 5(3), 1-28.

45 Yan, F., Eilers, M., Lüdtke, A., & Baumann, M. (2016, June). Developing a model of driver's uncertainty in lane change situations for trustworthy lane change decision aid systems. In 2016 IEEE Intelligent Vehicles Symposium (IV) (pp. 406-411). IEEE.

46 Zhang, Y., Wu, C., Qiao, C., Sadek, A., & Hulme, K. F. (2022). A Cognitive Computational Model of Driver Warning Response Performance in Connected Vehicle Systems. IEEE Transactions on Intelligent Transportation Systems.

47 Abbink, D. A., Mulder, M., & Boer, E. R. (2012). Haptic shared control: smoothly shifting control authority?. Cognition, Technology & Work, 14(1), 19-28.

48 Bärgman, J., Boda, C. N., & Dozza, M. (2017). Counterfactual simulations applied to SHRP2 crashes: The effect of driver behavior models on safety benefit estimations of intelligent safety systems. Accident Analysis & Prevention, 102, 165-180.

49 Bianchi Piccinini, G., Lehtonen, E., Forcolin, F., Engström, J., Albers, D., Markkula, G., ... & Sandin, J. (2020). How do drivers respond to silent automation failures? Driving simulator study and comparison of computational driver braking models. Human factors, 62(7), 1212-1229.

50 Colombi, J. M., Miller, M. E., Schneider, M., McGrogan, M. J., Long, C. D. S., & Plaga, J. (2012). Predictive mental workload modeling for semiautonomous system design: Implications for systems of systems. Systems Engineering, 15(4), 448-460.

51 Gunzelmann, G., Moore Jr, L. R., Salvucci, D. D., & Gluck, K. A. (2011). Sleep loss and driver performance: Quantitative predictions with zero free parameters. Cognitive Systems Research, 12(2), 154-163.

52 He, D., Wang, Z., Khalil, E. B., Donmez, B., Qiao, G., & Kumar, S. (in press). Classification of driver cognitive load: Exploring the benefits of fusing eye-tracking and physiological measures. Transportation Research Record.

53 Janssen, C. P., & Brumby, D. P. (2010). Strategic adaptation to performance objectives in a dual-task setting. Cognitive science, 34(8), 1548-1560.

54 Janssen, C. P., Brumby, D. P., & Garnett, R. (2012). Natural break points: The influence of priorities and cognitive and motor cues on dual-task interleaving. Journal of Cognitive Engineering and Decision Making, 6(1), 5-29.

55   Janssen, C. P., Boyle, L. N., Kun, A. L., Ju, W., & Chuang, L. L. (2019). A hidden markov framework to capture human–machine interaction in automated vehicles. International Journal of Human–Computer Interaction, 35(11), 947-955.

56   Janssen, C. P., Boyle, L. N., Ju, W., Riener, A., & Alvarez, I. (2020). Agents, environments, scenarios: A framework for examining models and simulations of human-vehicle interaction. Transportation research interdisciplinary perspectives, 8, 100214.

57   Janssen, C. P., Everaert, E., Hendriksen, H. M., Mensing, G. L., Tigchelaar, L. J., & Nunner, H. (2019). The influence of rewards on (sub-) optimal interleaving. PloS one, 14(3), e0214027.

58   Janssen, C. P., Iqbal, S. T., Kun, A. L., & Donker, S. F. (2019). Interrupted by my car? Implications of interruption and interleaving research for automated vehicles. International Journal of Human-Computer Studies, 130, 221-233.

59   Johora, F. T., & Müller, J. P. (2018, November). Modeling interactions of multimodal road users in shared spaces. In 2018 21st International Conference on Intelligent Transportation Systems (ITSC) (pp. 3568-3574). IEEE.

60   Kahl, S., Wiese, S., Russwinkel, N., & Kopp, S. (2021). Towards autonomous artificial agents with an active self: modelling sense of control in situated action. Cognitive Systems Research, 72, 50-62. https://doi.org/10.1016/j.cogsys.2021.11.005

61   Khosroshahi, E. B., Salvucci, D. D., Veksler, B. Z., & Gunzelmann, G. (2016). Capturing the effects of moderate fatigue on driver performance. In Proceedings of the 14th International Conference on Cognitive Modeling (pp. 163-168).

62   Lotz, A., Russwinkel, N., Wagner, T. and Wohlfarth, E. (2020). An adaptive assistance system for subjective critical driving simulation: understanding the relationship between subjective and objective complexity. In D. de Waard et al. (Eds.), Proceedings of the Human Factors and Ergonomics Society Europe Chapter 2019 Annual Conference. Nantes, France, p. 97-108.

63   Lotz, A., Wiese, S. & Russwinkel, N. (2019). SEEV-VM: ACT-R Visual Module based on SEEV theory. In Proceedings of the 17th International Conference on Cognitive Modelling (ICCM 2019), Montreal, Canada (pp. 301-306).

64   Markkula, G., Romano, R., Madigan, R., Fox, C. W., Giles, O. T., & Merat, N. (2018). Models of human decision-making as tools for estimating and optimizing impacts of vehicle automation. Transportation research record, 2672(37), 153-163.

65   Mole, C., Pekkanen, J., Sheppard, W., Louw, T., Romano, R., Merat, N., ... & Wilkie, R. (2020). Predicting takeover response to silent automated vehicle failures. Plos one, 15(11), e0242825.

66   Nagaraju, D., Ansah, A., Ch, N. A. N., Mills, C., Janssen, C. P., Shaer, O., & Kun, A. L. (2021). How Will Drivers Take Back Control in Automated Vehicles? A Driving Simulator Test of an Interleaving Framework. In 13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (pp. 20-27).

67   Pekkanen, J., Giles, O. T., Lee, Y. M., Madigan, R., Daimon, T., Merat, N., & Markkula, G. (2021). Variable-drift diffusion models of pedestrian road-crossing decisions. Computational Brain & Behavior, 1-21.

68   Salvucci, D. D. (2006). Modeling driver behavior in a cognitive architecture. Human factors, 48(2), 362-380.

69   Salvucci, D. D., & Gray, R. (2004). A two-point visual control model of steering. Perception, 33(10), 1233-1248.

70   Salvucci, D. D., Zuber, M., Beregovaia, E., & Markley, D. (2005). Distract-R: Rapid prototyping and evaluation of in-vehicle interfaces. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 581-589).

71 Schwall, M., Daniel, T., Victor, T., Favaro, F., & Hohnhold, H. (2020). Waymo public road safety performance data. arXiv preprint arXiv:2011.00038.

72 Svärd, M., Markkula, G., Bärgman, J., & Victor, T. (2021). Computational modeling of driver pre-crash brake response, with and without off-road glances: Parameterization using real-world crashes and near-crashes. Accident Analysis & Prevention, 163, 106433.

73 Trende A., Unni A., Rieger J., Fraenzle M. (2021) Modelling Turning Intention in Unsignalized Intersections with Bayesian Networks. In: Stephanidis C., Antona M., Ntoa S. (eds) HCI International 2021 – Posters. HCII 2021. Communications in Computer and Information Science, vol 1421. Springer, Cham.

74 van Maanen, L., Heiden, R. V. D., Bootsma, S., & Janssen, C. P. (2021). Identifiability and Specificity of the Two-Point Visual Control Model of Steering. In Proceedings of the Annual Meeting of the Cognitive Science Society (Vol. 43, No. 43).

75 Wickens, C. D. (2015). Noticing events in the visual workplace: The SEEV and NSEEV models. In: R. R. Hoffman, et al. eds. Part VI – Perception and Domains of Work and Professional Practice. Cambridge: Cambridge University Press, pp. 749-768.

76 Wiese, S., Lotz, A., & Russwinkel, N. (2019). Seev-vm: Act-r visual module based on seev theory. In Proceedings of the 17th International Conference on Cognitive Modeling (pp. 301-307).

6.2 Research agenda to further the field

Christian P. Janssen (Utrecht University, NL), Martin Baumann (Universität Ulm, DE), Shamsi Tamara Iqbal (Microsoft – Redmond, US), and Antti Oulasvirta (Aalto University, FI)

During the conference all the attendees identified various areas that can be part of a longer-term research agenda to move the field forward. Below, we list some of these research agenda items very briefly. They are clustered within each of the broader challenges that were a core panel discussion. Of course, these challenges are only an incomplete subset of the many research questions that are out there.

Challenge 1: What phenomena and driving scenarios need to be captured?

- Trust: For example, models could be used to study when is there (too much of) a deviation from an "appropriate level of trust"? But the deeper question is How to model / predict this in an online way?
- Mental models: How can models be used to aid people to have a good (more realistic, detailed) mental model of car's capabilities and feedback? And also vice-versa: how can the can have an appropriate mental model of the driver (and their knowledge).

Challenge 2: What technical capabilities do computational models possess?

- In what areas can models be part of the driving system and actually be used on the road?
- What are the requirements and constraints from a legal perspective when models are used on the road?
- Different models have different goals. There has been discussion whether we need computational models to model actual driving (i.e., to be able to take control of a vehicle, or be a "digital twin" of the driver of a vehicle), or whether models are mostly used for understanding human behavior and during the design phases?
- A related question is then: Where can or should a driving model be a black box or white box model?

Challenge 3: How can models benefit from advances in AI while avoiding pitfalls?

- All models have some mechanisms that link inputs to outputs. However, the components come from different places: theory ("white box"), data ("black box"), or a combination? Each technique has advantages and disadvantages. How can techniques and insights from white box and black box models be best combined?
- This ties in with a more general challenge of how to best balance between generalizability /variability and fixed structure of a model. What is truly fixed (and well represented in a model)? What is learned / variable? How can one know they have a properly generalizable model?
- Crafting a model that has both white box and black box items can also be seen as a scientific process or method in itself. This method can be more standardized.
- Researchers make quite some choices during the model selection (and developing) process. How can this be approached in a more principled manner?

Challenge 4:What insights are needed for and from empirical research?

- For what problems are large datasets needed? And for what problems are small datasets sufficient? (i.e., balancing also more traditional statistical techniques with machine learning techniques)
- What is needed to then make correct inferences on both? Small datasets and big datasets each have their charms, but require different techniques and insights. Just because a dataset is larger, does not make its quality better nor does it mean that the inferences are more reliable (as some form of data quality needs to be ensured).
- Some form of standardized data set might be useful for model development and model competition. A research objective can be to develop and grow such a data set (and let it grow over time).
- The above is tied to a need to have benchmark tests / phenomena to test models on. Is there perhaps a "golden standard" test?

Challenge 5: How can models inform design and governmental policy?

◧ How can cognitive models inform (the design of) future interaction best? Design efforts often explore specific scenarios, but within that look at various alternative designs. By comparisons, models sometimes are more fixed towards specific methods or outcomes. How can they also incorporate that flexibility?

◧ And how can the appropriate (modelling / model sketching) tools be made?

Participants

- Martin Baumann
  Universität Ulm, DE
- Jelmer Borst
  University of Groningen, NL
- Alexandra Bremers
  Cornell Tech – New York, US
- Duncan Brumby
  University College London, GB
- Debargha Dey
  TU Eindhoven, NL
- Patrick Ebel
  Universität Köln, DE
- Martin Fränzle
  Universität Oldenburg, DE

- Luisa Heinrich
  Universität Ulm, DE
- Moritz Held
  Universität Oldenburg, DE
- Jussi Jokinen
  University of Jyväskylä, FI
- Dietrich Manstetten
  Robert Bosch GmbH –
  Stuttgart, DE
- Gustav Markkula
  University of Leeds, GB
- Roderick Murray-Smith
  University of Glasgow, GB

- Antti Oulasvirta
  Aalto University, FI
- Nele Rußwinkel
  TU Berlin, DE
- Shadan Sadeghian Borojeni
  Universität Siegen, DE
- Hatice Sahin
  Universität Oldenburg, DE
- Philipp Wintersberger
  TU Wien, AT
- Fei Yan
  Universität Ulm, DE

Remote Participants

- Linda Ng Boyle
  University of Washington –
  Seattle, US
- Lewis Chuang
  LMU München, DE
- Benjamin Cowan
  University College – Dublin, IE
- Birsen Donmez
  University of Toronto, CA
- Justin Edwards
  ADAPT Centre – Dublin, IE
- Mark Eilers
  Humatects – Oldenburg, DE
- Shamsi Tamara Iqbal
  Microsoft – Redmond, US

- Christian P. Janssen
  Utrecht University, NL
- Myounghoon Jeon
  Virginia Polytechnic Institute –
  Blacksburg, US
- Xiaobei Jiang
  Beijing Institute of
  Technology, CN
- Wendy Ju
  Cornell Tech – New York, US
- Tuomo Kujala
  University of Jyväskylä, FI
- Andrew Kun
  University of New Hampshire –
  Durham, US

- Otto Lappi
  University of Helsinki, FI
- Nikolas Martelaro
  Carnegie Mellon University –
  Pittsburgh, US
- Andreas Riener
  TH Ingolstadt, DE
- Boris van Waterschoot
  Rijkswaterstaat – Utrecht, NL
- Yiqi Zhang
  Pennsylvania State University –
  University Park, US