# Safe Handover in Mixed-Initiative Control for Cyber-Physical Systems

**Frederik Wiehr**
**Anke Hirsch**
**Florian Daiber**
**Antonio Krüger**
German Research Center for Artificial Intelligence (DFKI)
Saarland Informatics Campus

**Alisa Kovtunova**
**Stefan Borgwardt**
Institute of Theoretical Computer Science, TU Dresden
stefan.borgwardt@tu-dresden.de

**Ernie Chang**
**Vera Demberg**
Saarland University
Saarland Informatics Campus
vera@coli.uni-saarland.de

**Marcel Steinmetz**
**Jörg Hoffmann**
Saarland University
Saarland Informatics Campus
hoffmann@cs.uni-saarland.de

*CHI'20*, Designing Safety Critical Interactions: Hunting Down Human Error" CHI 2020 Workshop, April 25, 2020, Honolulu, HI, USA

## Abstract

For mixed-initiative control between cyber-physical systems (CPS) and its users, it is still an open question how machines can safely hand over control to humans. In this work, we propose a concept to provide technological support that uses formal methods from AI – description logic (DL) and automated planning – to predict more reliably when a handover is necessary, and to increase the advance notice for handovers by planning ahead of runtime. We combine this with methods from human-computer interaction (HCI) and natural language generation (NLG) to develop solutions for safe and smooth handovers and provide an example autonomous driving scenario. A study design is proposed with the assessment of qualitative feedback, cognitive load and trust in automation.

## Author Keywords

criticality; mixed-initiative control; safety; cyber-physical system; car driving; multimodal interaction; human-computer interaction; natural language generation; description logics; automated planning; critical interaction; interdisciplinary

## CCS Concepts

•**Human-centered computing** → **Human computer interaction (HCI);** User studies;

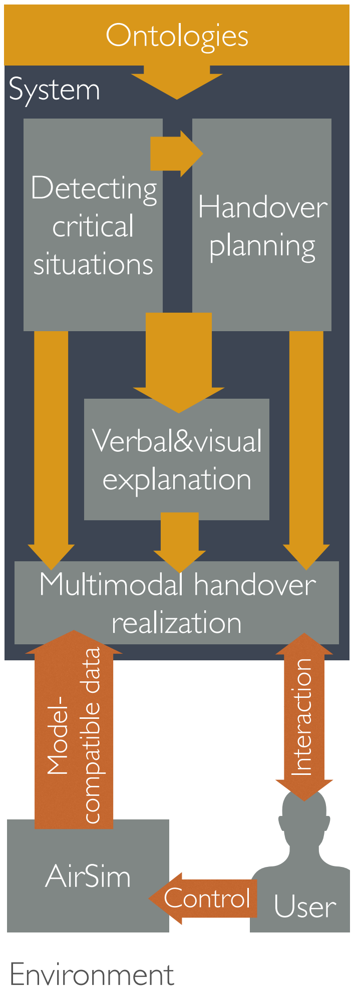


**Figure 1:** Proposed system architecture

## Introduction

Mixed-initiative control systems have shown that when decisions are made or suggested by automated systems it is essential that an explanation should be provided [10]. An example of such a system is found in autonomous driving, where the cyber-physical system (CPS) derives responses from navigation functions based on human input.

The purpose of this paper is to develop a framework to support a safe and stress-limiting handover of control to the driver, combining benefits derived from formal AI methods, human-computer interaction (HCI) and natural language generation (NLG). The system must alert the user in the "best possible way" when a handover is required. Practically, this poses the challenge that human factors may render the handover process difficult to achieve [8], in part due to the lack of *situation awareness* in the presence of a secondary task embedded in the experimental setups. The driver, if distracted, must be able to promptly re-establish awareness of the situation, using sensory cues from the environment. The autonomous car can support this situation awareness by communicating its knowledge of the situation at the time of the transition. Moreover, the system may interact with the user, or takeover in the case of non-response.

Towards this end, we describe our implemented framework which utilizes the relevant environment-state information from the open-source simulator for the autonomous system AirSim [28], and propose a preliminary study design on the effect of varying the modality and timeliness of explanations in a simulated driving experience, with four handover situations.

## Related Work

The related work in HCI can be separated into user modeling, situation awareness/vigilance, and multimodal interaction.

*User modeling* enables a system to maintain a conceptual understanding of the user (user model [19], in which user differences need to be modeled explicitly). Tailoring handover requests by the system to individual users makes it possible to take into account the user's experience as well as individual differences relating to cognitive capacity.

*Situation awareness* describes the human's awareness of the environment, e.g. a critical situation for a task at hand [8]. Such information depends strongly on the situation (e.g. a pilot approaching an airport vs. a driver navigating in dense traffic). Prior research utilizes operational quantitative human factors to assess the situation awareness of human operators [1]. In highly automated systems, the risk of humans being out-of-the-loop increases and thus a potential handover is more difficult to achieve [8]. The degree of *vigilance* influences the ability of a human to attend to the environment. Past research has developed vigilance measures based on questionnaires, and sensors assessing heart rate, eye movement and skin conductance [7].

*Multimodal interaction* has the potential to increase the usability and thus the safety of operation [3]. It has been used for mobile applications and environments, including gesture and speech [29], eye tracking and face detection [23], and tangible interaction [18] to adapt to the user's needs. Moreover, the styles of visualization can improve "trust" in autonomous driving [13], while identified feedback factors [20, 13] improve the understandability and trust of system decisions made in autonomous driving.

## Architecture

*Planning*

The reasoning capabilities of modern AI planners [9, 27] provide the possibility to foresee *critical situations* the autonomous system's decisions may lead to in the future, i.e. situations that cannot be reliably handled by the system on its own. Identifying such situations ahead of time is not only crucial for a successful transfer of control to the user, but may actually allow the handover to be avoided altogether in certain cases.

To identify and to anticipate critical situations, a planner builds upon an abstract model consisting of two major components: (1) *state features* that allow representation of an abstract view of the world for a specific point in time (e.g. current position and speed of the car), and the (2) *actions* the autonomous system can do at an abstract level (e.g. accelerating, changing lanes).

In our architecture, AI planning is used for two purposes: (1) *monitoring*, and (2) *replanning*. By default, in (1) the planner is only used to test the autonomous system's decisions via simulations within the abstract world model; and (2) uses planning to check the existence of an alternative to the autonomous system's decisions in order to avoid other critical situations.

*Description Logics*

Knowledge representation based on description logics (DLs) allows us to describe the complex environment in a so-called ontology, specifying constraints for the system states and for reasoning about the domain knowledge. Ontologies have, for example, been proposed for real-time patient monitoring [17, 26], detecting composite dance movements in annotated ballet videos [24], and weather and turbine monitoring [2].

By design, description logics are close to human reasoning and can supply explanations for decisions made by a cyber-physical system based on an ontology. The DL component in our architecture provides input for the planning component. By performing Boolean temporal query answering over a sequence of potential future states, the DL component assesses the criticality level of this route. Temporal queries describe different potentially dangerous situations on a road. The level of criticality that is reported to the planning component depends on the number and severity of the situations that are detected.

*Natural Language Generation*

Description logic expressions serve as the semantic inputs to natural language generation (NLG) systems, which is analogous to the task of data-to-text generation. Traditionally, this is dealt with using a pipeline consisting of *content planning*, *sentence planning* and *linguistic realization* [25]. While the format of the data varies from task to task, it typically involves the linearization step [22] where structured data are converted into sequences, before being processed by downstream systems for *linguistic realization*.

In the task of delivering handover message, NLG systems face a two-fold challenge: (1) there is insufficient annotated data for every situation, (2) handover messages should not cognitively overload the user, but instead be at a situationally-appropriate level; see [6, 4, 12]. To this end, our proposed NLG system adopts (A) semi-supervised NLG techniques [16] that minimize the required data and (B) an automatic quality estimator [5] that assesses the information density of the message.

*Human-Computer Interaction*

For safe handovers in this context, the interaction side needs to be considered and adapted depending on the anticipated critical situation by the planner and DL component.

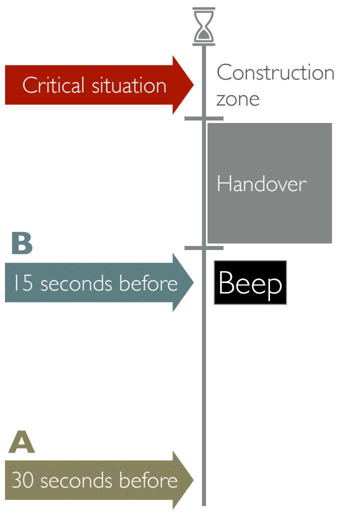


**Figure 2:** Handover Timing

Therefore, user modeling techniques can be applied to assess differences implied by the situation (such as cognitive load and vigilance: see below), as well as slowly changing individual user differences such as task familiarity (e.g. differentiating between novices and expert operators) [14]. To provide a basis for adaptations such as a multimodal handover realization, the HCI component of the system uses ontologies to provide access to and allow for interpretation of the user models [15]. Although general user modeling is already established, the special safety-critical aspects of handover situations are still insufficiently addressed by prior research. Interface adaptations for cyber-physical systems have to be carefully integrated with the handover planning and DL components with a focus on safety critical situations. This includes developing concepts that respect cognitive implications of human operators.

## Conclusion & Outlook

To gain insights into the use of a planner combined with description logics, we have chosen highly automated driving as a use case for handover in safety critical situations. Specifically, we plan to investigate handover scenarios where the driver has to take over control of the car in a driving simulator while performing a secondary task. The car or the systems can signal the handover in four different conditions. A goal here would be to test whether drivers would perceive additional verbal explanations as beneficial compared to a classical handover technique. The classic handover simply issues a notification at point B. Furthermore, the beep could be combined with either (a) a preceding request to take over (planner) or (b) a subsequent explanation about why the user had to take over (DL). The last condition could be a combination of planner and DL where the participants get an explanation and a request before the beep which signals them to take over. Planning enables the car to initiate the handover in time (at point A, see Figure 2) whereas the description logic alone can only generate the explanation in situ (at B). The criticality of the handover situation can be increased by approaching a construction zone in the driving simulator. Furthermore, the cause of an additional danger can be introduced by different driving scenarios of the cars in front of the participant. To avoid learning effects, four different driving scenarios should be combined with the four aforementioned conditions, like a unpredictable driver with odd steering behavior, a car that drives to the left lane, a very slow driver with sudden braking or a truck with an unsecured load. Quantitative as well as qualitative data will be collected to conduct statistical analysis as well as participants' personal evaluation of each handover condition. The NASA-TLX [11] and the Trust in Automation Questionnaire (TiA) [21] could assess cognitive load and trust in automated systems, respectively. Last, a semi-structured interview can be conducted to let the participants rank the different conditions and give reasons for the ranking.

A current design error in today's safety-critical systems is that these do not feature built-in concepts to pre-plan, or to recognize and explain problem causes to the user. Programs running in CPS participate in actions and decisions that affect humans, especially in highly automated vehicles, when a handover needs to be performed in critical situations. With the approach of combining formal methods with HCI, we believe that generating verbal and visual explanations in a timely manner using planning and description logics can ease the process for the user of regaining situational awareness and allow for a safe handover of control. We would like to further develop this idea during the workshop and discuss possible scenarios and experimental designs.

**Acknowledgments**

This work was partially supported by DFG grant 389792660 as part of TRR 248 (https://perspicuous-computing.science) and BMBF grant 01IW17004 (http://tractat.dfki.de).

# REFERENCES


[1] Marilyn Jager Adams, Yvette J. Tenney, and Richard W. Pew. 1995. Situation awareness and the cognitive management of complex systems. *Human Factors* 37, 1 (1995), 85–104. DOI: http://dx.doi.org/10.1518/001872095779049462

[2] Sebastian Brandt, Elem Güzel Kalayci, Roman Kontchakov, Vladislav Ryzhikov, Guohui Xiao, and Michael Zakharyaschev. 2017. Ontology-Based Data Access with a Horn Fragment of Metric Temporal Logic. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAIâĂŹ17)*. AAAI Press, 1070âĂŞ1076.

[3] Philip R. Cohen and David R. McGee. 2004. Tangible multimodal interfaces for safety-critical applications. *Commun. ACM* 47, 1 (2004), 41–46.

[4] Matthew W Crocker, Vera Demberg, and Elke Teich. 2016. Information density and linguistic encoding (IDeaL). *KI - Künstliche Intelligenz* 30, 1 (2016), 77–81.

[5] José GC de Souza, Michael Kozielski, Prashant Mathur, Ernie Chang, Marco Guerini, Matteo Negri, Marco Turchi, and Evgeny Matusov. 2018. Generating e-commerce product titles and predicting their quality. In *Proceedings of the 11th International Conference on Natural Language Generation*. 233–243.

[6] Vera Demberg and Asad Sayeed. 2011. Linguistic cognitive load: Implications for automotive UIs. *Proc. of AutomotiveUI* 2011 (2011).

[7] Mica R. Endsley. 1995. Measurement of situation awareness in dynamic systems. *Human Factors* 37, 1 (1995), 65–84.

[8] Mica R. Endsley. 1996. Automation and situation awareness. *Automation and human performance: Theory and applications* (1996), 163–181.

[9] Malik Ghallab, Dana S. Nau, and Paolo Traverso. 2004. *Automated planning - theory and practice*. Elsevier.

[10] Bryce Goodman and Seth R. Flaxman. 2017. European Union Regulations on Algorithmic Decision-Making and a "Right to Explanation". *AI Magazine* 38, 3 (2017), 50–57. https://www.aaai.org/ojs/index.php/aimagazine/article/view/2741

[11] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, 139–183.

[12] Katja I. Häuser, Vera Demberg, and Jutta Kray. 2019. Effects of Aging and Dual-Task Demands on the Comprehension of Less Expected Sentence Continuations: Evidence From Pupillometry. *Frontiers in Psychology* 10 (March 2019). DOI: http://dx.doi.org/10.3389/fpsyg.2019.00709

[13] Renate Häuslschmid, Max von Bülow, Bastian Pfleging, and Andreas Butz. 2017. SupportingTrust in Autonomous Driving. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces (IUI âĂŹ17)*. Association for Computing Machinery, New York, NY, USA, 319âĂŞ329. DOI: http://dx.doi.org/10.1145/3025171.3025198

[14] Dominik Heckmann and Antonio Krüger. 2003. A user modeling markup language (UserML) for ubiquitous computing. In *International Conference on User Modeling (UM)*. 393–397.

[15] Dominik Heckmann, Tim Schwartz, Boris Brandherrn, Michael Schmitz, and Margeritta von Wilamowitz-Moellendorff. 2005. Gumo – the general user model ontology. In *International Conference on User Modeling (UM)*. 428–432.

[16] Xudong Hong, Ernie Chang, and Vera Demberg. 2019. Improving Language Generation from Feature-Rich Tree-Structured Data with Relational Graph Convolutional Encoders. In *Proceedings of the 2nd Workshop on Multilingual Surface Realisation (MSR 2019)*. 75–80.

[17] Anna Hristoskova, Vangelis Sakkalis, Giorgos Zacharioudakis, Manolis Tsiknakis, and Filip De Turck. 2014. Ontology-driven monitoring of patient's vital signs enabling personalized medical detection and alert. *Sensors* 14, 1 (2014), 1598–1628. DOI: http://dx.doi.org/10.3390/s140101598

[18] Vaiva Kalnikaite, Stefan Kreitmayer, Yvonne Rogers, Jon Bird, Nicolas Villar, Khaled Bachour, Stephen Payne, Peter M. Todd, Johannes Schöning, and Antonio Krüger. 2011. How to nudge in Situ. In *Proceedings of the 13th international conference on Ubiquitous computing - UbiComp '11*. ACM Press. DOI: http://dx.doi.org/10.1145/2030112.2030115

[19] Alfred Kobsa and Wolfgang Wahlster. 1989. *User Models in Dialog Systems*. Springer.

[20] Jeamin Koo, Jungsuk Kwac, Wendy Ju, Martin Steinert, Larry Leifer, and Clifford Nass. 2015. Why did my car just do that? Explaining semi-autonomous driving actions to improve driver understanding, trust, and performance. *International Journal on Interactive Design and Manufacturing (IJIDeM)* 9, 4 (2015), 269–275.

[21] Moritz Körber. 2018. Theoretical considerations and development of a questionnaire to measure trust in automation. In *Congress of the International Ergonomics Association*. Springer, 13–30.

[22] Simon Mille, Anja Belz, Bernd Bohnet, Emily Pitler, and Leo Wanner (Eds.). 2018. *Proceedings of the First Workshop on Multilingual Surface Realisation*. Association for Computational Linguistics, Melbourne, Australia. https://www.aclweb.org/anthology/W18-3600

[23] Jörg Müller, Juliane Exeler, Markus Buzeck, and Antonio Krüger. 2009. Reflective Signs: Digital signs that adapt to audience attention. In *IEEE International Conference on Pervasive Computing and Communications (PerCom)*. 17–24.

[24] Katerina El Raheb, Theofilos Mailis, Vladislav Ryzhikov, Nicolas Papapetrou, and Yannis E. Ioannidis. 2017. BalOnSe: Temporal Aspects of Dance Movement and Its Ontological Representation. In *The Semantic Web – 14th International Conference, ESWC 2017, Portorož, Slovenia, May 28 - June 1, 2017, Proceedings, Part II*. 49–64. DOI: http://dx.doi.org/10.1007/978-3-319-58451-5_4

[25] Ehud Reiter and Robert Dale. 1997. Building applied natural language generation systems. *Natural Language Engineering* 3, 1 (March 1997), 57–87. DOI: http://dx.doi.org/10.1017/s1351324997001502

[26] Ahlem Rhayem, Mohamed Ben Ahmed Mhiri, Mayssa Ben Salah, and Faïez Gargouri. 2017. Ontology-based system for patient monitoring with connected objects. In *Knowledge-Based and Intelligent Information & Engineering Systems: Proceedings of the 21st International Conference KES-2017, Marseille, France, 6 - 8 September 2017*. 683–692. DOI: `http://dx.doi.org/10.1016/j.procs.2017.08.127`

[27] Silvia Richter and Matthias Westphal. 2010. The LAMA Planner: Guiding cost-based anytime planning with landmarks. *J. Artif. Intell. Res.* 39 (2010), 127–177. DOI: `http://dx.doi.org/10.1613/jair.2972`

[28] Shital Shah, Debadeepta Dey, Chris Lovett, and Ashish Kapoor. 2018. AirSim: High-fidelity visual and physical simulation for autonomous vehicles. In *Field and Service Robotics*. Springer, 621–635.

[29] Rainer Wasinger, Antonio Krüger, and Oliver Jacobs. 2005. Integrating intra and extra gestures into a mobile and multimodal shopping assistant. In *International Conference on Pervasive Computing*. 297–314.