A Cognitive Model of Drivers Attention

Kerstin Sophie Haring (ksharing@fennel.rcast.u-tokyo.ac.jp)
Research Center for Advanced Science and Technology, The University of Tokyo
4-6-1, Komaba, Meguro-ku, Tokyo, 153-8904, Japan

Marco Ragni (ragni@cognition.uni-freiburg.de)
Lars Konieczny (lars@cognition.uni-freiburg.de)
Center for Cognitive Science, University of Freiburg
Friedrichstr. 50, 79098 Freiburg, Germany

Abstract

Cognitive architectures can account for highly complex tasks. One of the greatest challenges is understanding and modeling human driving behavior. This paper describes an integrated cognitive model of human attention during the performance of car driving. In this task, the attention process can be divided into at least three basic components: the control process, the monitoring process, and finally, the decision making process. Of these basic tasks, the first has the highest priority. All three phases are implemented in a cognitive model in the cognitive Architecture ACT-R 6.0. The model is able to keep a traffic lane, overtake another vehicle by lane change, identifies traffic signs and different situations emerging at crossroads.

Keywords: Driver behavior model; cognitive architecture; ACT-R; Attention

Introduction

Even for long-time practitioners driving a car is a highly complex task. This becomes evident by the still high number of accidents. E.g., in 2010 in Germany nearly 375.000 persons were injured in approximately 290.000 automobile accidents (Statistisches Bundesamt, 2011). In about 84% of all cases the cause of an accident could be traced back to driver errors (cp. Fig. 1). Nowadays passive safety systems like the airbag are reaching their technological limits and the focus shifts more towards active safety systems. Active systems, however, require exact knowledge about the driver, the vehicle, and the environment. To increase the acceptance of active intervention through the safety systems in cars, these systems should act in accordance to the driver. The driver and the human driving behavior must be considered for the future development of safety systems. Consequently, one focus of research is to analyze human behavior and predict possible errors.

We present the implementation of a cognitive driver model, simulating human attention and driving behavior. A driver model can be a powerful instrument with several possible fields of application, such as the development of intelligent driver assistant systems. The model is an adaption of Salvuccis`s (2006) driver model developed in the Cognitive Architecture ACT-R 5. Our model is implemented the newer version ACT-R 6 (Anderson, 2007) and using the standard ACT-R development environment running on an open source LISP, which not only guarantees support and accountability, but also enables the research community to use the developed model for further research. It is able to keep a traffic lane, initiate and decide about a change of the lane in case of upfront traffic, identify prevalent situations at crossroads and react to traffic signs.

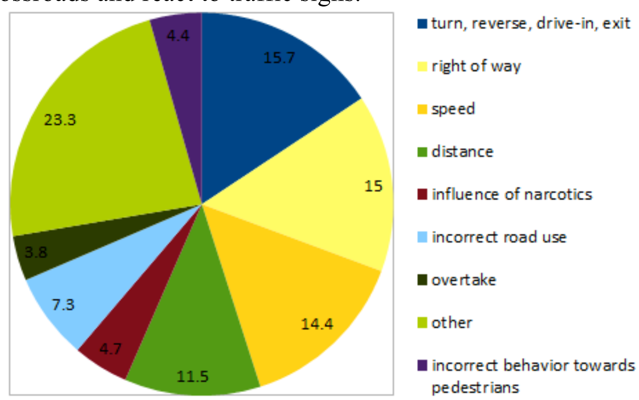

Fig. 1: Driver errors in automobile accidents with person injury (Statistisches Bundesamt, 2011).

Previous work

Most developed approaches can be distinguished into two classes: task specific and generic approaches. *Task specific approaches* such as Cosmodrive (Bellet et al., 2007) and Pelops (Benmimoun, 2004) reproduce the cognitive functions of a car driver. In contrast to task specific approaches, *generic approaches* can model various aspects of human behavior. Therefore, it is necessary for these architectures to include a theory of human information processing. Examples for such architectures in which driver models have been implemented are ACT-R (Anderson, 1993; Salvucci, 2006), SOAR (Aasman,1995) and QN-MHP (Liu et al., 2006).

Previous models can be divided into three categories: First, early models concentrated mainly on steering and lane keeping. These models focus on the control process and are able to detect some cognitive aspects, but according to Boer (1999) they are highly dependent on difficult perceivable inputs from the environments. Second category comprises perception-action models which are through the perceptual constraints oriented closer on human behavior (e.g. Rushton et al., 1998; Salvucci & Gray, 2004; Wilkie & Wann, 2003). Yet, these models do not allow for movement dynamics.

Finally, the third category includes models that are trying to unify the various aspects of a driving task and are therefore the most closely associated to the here presented work. These models not only explore and unify the various aspects of driving behavior, they also explore the generality of the cognitive architectures used for their development. Driver models were described by Aasman (1995) in the cognitive architecture SOAR and by Liu (1996) in Queuing Network-Model Human Processor (QN-MHP). Although these models already exist in other cognitive architectures and the central ideas remain the same in any architecture, the ACT-R model of a driver shows a broader spectrum of application (Salvucci 2001; 2006).

Salvucci (2006) developed a first integrated cognitive model of human driving behavior in ACT-R. He showed in his work the generality and the applicability using the cognitive architecture ACT-R for the specific task of driving. His model is designed to keep a standard vehicle on a multi-lane highway with moderate traffic. The model is also able to recognize the distance to a vehicle ahead and to make the decision for overtaking. As driving is a highly complex task and not readily implementable, this model has some limitations. The model solely was meant to interact with a highway environment without recognition of traffic signs, crossings or slip roads. An implementation limitation was the use of the previous version ACT-R 5.0 and its incompatibility to newer versions. It was also not possible to make the ACT-R model interact directly with a driving simulator.

The cognitive architecture

A cognitive architecture compromises theories about the operation mode of human information processing and aims at using procedures similar to humans. In other words, it describes a comprehensive computer model of human cognition. ACT-R (Anderson, 1993; Anderson 2007) is such a comprehensive theory of human cognitive capacities. It is also a modeling environment, used to describe human cognitive processes. Most of its basic assumptions are inspired by the progress of cognitive neuroscience. ACT-R is a framework in which the researcher can create models (programs) for different tasks. Running this model produces a simulation of human behavior. The main assumption of ACT-R is the representation of knowledge as either declarative or procedural knowledge. Declarative knowledge, consisting of facts, is represented in form of *chunks,* or small logical units which encode simple facts (e.g. the fact: "Berlin is in Germany"). Procedural knowledge, representing knowledge about how we do things, is represented in form of *production rules,* condition-action rules that generate a specific action (e.g. manipulate declarative knowledge) if the conditions of this rule are fulfilled.

In other words, ACT-R's knowledge representation is split in two kind of memory modules. Modules can be accessed through their buffers. The state of ACT-R at a given time is the content of the buffers at that time. Buffers are connected to the modules and are changed by production rules. Every buffer and (nearly) every module can be allocated to a cortex region. This enables an interesting mapping between buffers and neural processes (Anderson 2007).

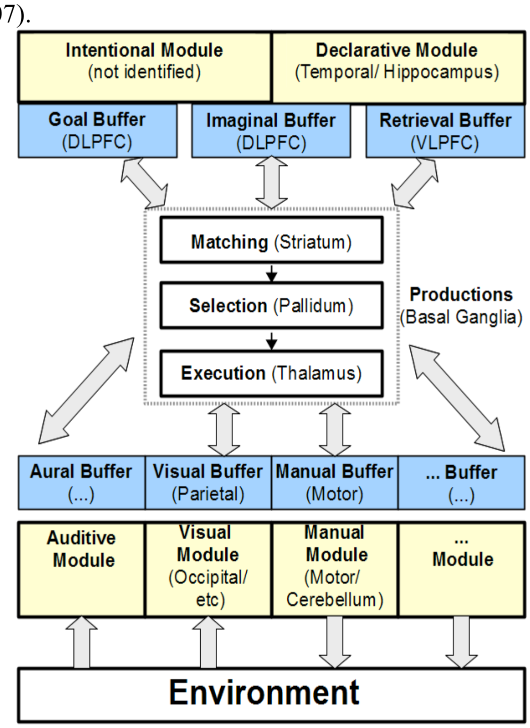

Fig. 2: The organization of information the cognitive architecture ACT-R (Anderson, 1993). The buffers contain information and are connected to modules associated with brain regions.

Cognitive model

We introduce now a computational model of human attention in a car driving task implemented in the ACT-R architecture. It models human attention and behavior for driving a car on a straight road, overtaking another vehicle by lane-change, identifying a traffic sign and crossroads.

Driver Modeling

The goal of this research was to develop an integrated driver model in the context of embodied cognition, task and artifact (ETA) framework. Byrne (2001) describes the ETA framework as understanding of interactive behavior based on the Cognition-Task-Artifact triad introduced by Gray (Gray & Altman, 2001). Interactive behavior is a function of the performed Task, the Artifact (instrument) by which the task is performed, and the Embodied Cognition, the cognitive, perceptual and motor capabilities by which a person acts through the artifact.

Cognitive modeling of human driving behavior should address all three components. An integrated model considers the driving related tasks (Task), the interface between the human and the vehicle (Artifact) and the processes that execute the driving task on the vehicle (Embodied Cognition). The system must be specified regarding a detailed description of the artifact being used and the task to perform. Some successfully implemented and applied models only emphasize one or two of these components like the perception-action models of control of Fajen (Fajen & Warren, 2003), which provides a compact description of the behavioral dynamics of steering and obstacle avoidance, control-theoretic models like Donges (1978), dividing the steering task into a guidance and a stabilization level or machine-learning models, supporting automobile drivers steering by sampling an image, assessing the road curvature, and determining the lateral offset of the vehicle (Pomerleau & Jochem, 1996).

Driving is a continuously changing task of basic subtasks. These must be integrated and interleaved. This model uses three basic components, control, monitoring, and decision making (see Fig. 3), derived from the hierarchical control structure of Michon (1985). Michon identified three levels of skills and control for the driving task: operational (control), tactical (maneuvering), and strategic al (planning). He claims that a comprehensive model should take into account the various levels and also provide an information flow control that allows to switch from one level to the other.

The independent subtasks of a simple driving task (see Fig. 3) were implemented as *control*, the operational process controlling the input, *monitoring*, the tactical process interacting with the environment, and *decision making*, also analogous to the tactical level of Michon (1985), managing maneuvers like overtaking. These subtasks are processed serially. Every production of the top level goal *drive* has sub-goals, which incorporate the three components.

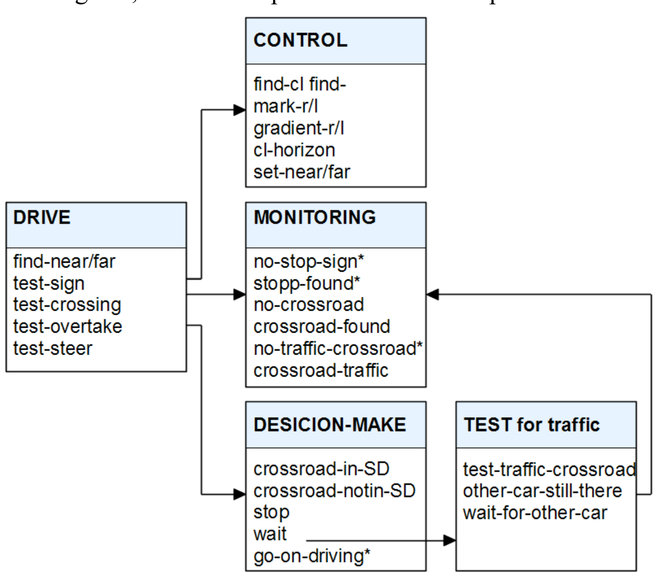

Fig. 3: Schematic representation of the production rules of the driver model in a simple crossroad scenario. The title of a box indicates the current goal and the corresponding production rules. The arrows show the flow of control and the asteriks the return to the parent-goal.

Development Environment The theory of ACT-R is embedded in the ACT-R software in form of Common Lisp functions. This model is implemented in Clozure CommonLisp 1.3 and the current version of ACT-R 6.0 under the operating system Ubuntu 9.04. In order to make the simulation environment interact with the ACT- R system, it was directly implemented in LISP with simple graphics and the extension with the LTK Lisp Toolkit. As it was not possible to make ACT-R directly interact with a driving simulator, we decided to use a Lisp-implementation of a driving environment.

Model Specification

As mentioned, the cognitive model of human attention integrated the three components control, monitoring and decision making. They are implemented as a loop of cognitive operations in the ACT-R serial processor.

The UML-Diagram in Fig. 4 shows the behavior of the cognitive model. To execute the task drive, the model runs through several states.

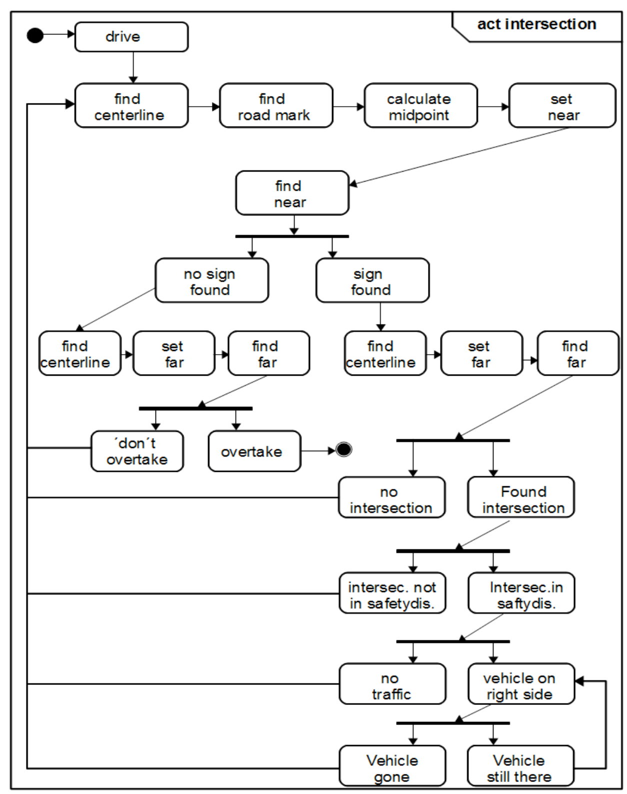

Fig. 4: UML-Diagram of the driver model

From the initial state, the model finds the road marks and sets the near point for stable navigation on the road. The model then fires a production rule screening for a traffic sign, changes the state according to the result and sets the far point. In our model, the near and far point are used as control components and explained in detail in the next paragraph. If the model reaches the state *find far* it can reach the state *overtake* or will repeat the control loop. If there is special state like an intersection, the model tests for other given constraints and according to the result of this test, is will either go to another special state or repeat the control loop updating the near and far points.

A crucial advantage of the ACT-R architecture is that the three components control, monitoring and decision-making can be implemented directly. This takes into account human constraints and results in a cognitive adequate model of human attention.

Control

The control component of attention while performing a driving task manages the perception of lower level visual cues and the control over the vehicle (e.g., stopping). The model uses the simple concept of two salient visual attributes. This concept is based on earlier findings on locomotion (Llewellyn, 1971) and steering. Further research (Donges, 1987; Land & Horwood, 1995) describes steering as divided in two levels, guidance and stabilization, by using a „far“ and a „near“ region. Models of steering developed under this assumption have been proven to be consistent with empirical evidence.

The perception of this model is based on the perception of two salient visual points (Salvucci & Gray, 2004), a near and a far point. These two points are used for guidance, stabilization and also, to observe other salient attributes. For the here created artificial road environment, these two points account to recognize relevant aspects in any situation which may arouse during driving a car.

The near point determines the position on the road, which is in the middle of the center line and the border line. To identify the direction of driving, the far point is used and usually set on the vanishing point on the horizon or on the lead vehicle. The far point is also used to identify other situations and can be set on non-control points like traffic signs or approaching cars. Fig. 5 illustrates the near and far points.

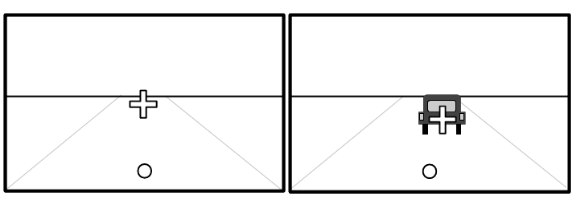

Fig. 5: Near and far points for a straight road with a vanishing point and a road segment with a lead car.

The ACT-R architecture limits the employment of the control component by using a serial cognitive processor. The serial processing of the subtasks is typical for the human bottleneck of information processing. The resulting model is not an optimal model in a mathematical sense, but approximates human behavior.

In a driving environment, the majority of lower level visual control is keeping the vehicle in the middle of the road lane, for which the near point is used. Although the far point is used to identify traffic signs, it mainly indicates the driving direction.

If the far point is not set on the vanishing point on the horizon, the model uses the combination of near and far point for determining the current scenario (see also Fig. 8 for an overview of implemented scenarios). If there is a lead vehicle, the distance between the two points is determined, and in case it falls below a certain safety distance, the model can react according to that (e.g. through slowing down or overtaking). In a crossroad scenario without an approaching car from the right hand side, the model will set the far point on the vanishing point of the horizon and continue driving. After that, this model will not look again for another car at the crossroad, which is surely an issue for future implementations. In case there is a vehicle or a stop sign, the stopping of the car is implemented here by setting the far point onto the near point. The model will continue a loop until the other vehicle is not on the crossroad anymore and out of the safety distance.

Monitoring

After the control component, the monitoring is one of the most important. Here, the environment is continuously captured (e.g. the model looks for a traffic signs) and updated in the declarative memory. In the here implemented driving environment, the situation awareness mainly focuses on other vehicles around, change of the scenario (from straight road to crossroad), or traffic signs. The model shifts the focus of visual attention towards a certain object which is then encoded as visual attribute. The shift could be based on a random-sampling model, checking the different environment areas with a probability *p,* which has been successfully done by Salvucci (2006). Here, the model monitors particular directions and visual attributes (e.g. other vehicles, center line) by an attention shift. The encoded attribute is noted in the declarative memory. As ACT-R has a build-in memory decay mechanism, it might be possible to predict driver errors because the chunks encoding the current environment decay and can be forgotten if not updated continuously. Another source of possible driver errors could be the potential failure in encoding relevant information (e.g. to overlook a traffic sign or a vehicle).

Decision Making

The information provided by the control and monitoring component is used to determine if and what decisions must be made on the tactical level concerning the maneuvering (e.g. stopping or overtaking). The most common decision making might be whether to stop or to continue driving. This decision depends on the traffic sign or on other vehicles. As described earlier, the execution of stopping corresponds simply to the use of the near and far points encoding current position and relevant aspects of the environment. In order for the model to produce a decision making process similar to humans, encoding a visual attribute and shifting visual attention cannot occur at the same time. For this model, the focus of attention is for example either on the near or far point or encoding a traffic sign. This restriction through serial processing seems to be a drawback in the sense of mathematical optimal behavior, but it describes the bottleneck typical for the human information processing (Anderson et al, 2004). Through the implementation of this restriction, it is possible to mimic human cognitive capacities, simulate the dynamic nature of human driving behavior, and therefore a cognitive adequate model of human driving behavior is produced.

The knowledge representation comprehends declarative knowledge in chunks and procedural knowledge in production rules. For example, the scenario at a crossroad was implemented in 73 explicit production rules, which are highly detailed and is therefore open to future extensions of the model. The control of attention in the ACT-R architecture is achieved through three different methods of shifting attention: First by specific locations or directions, second by specific characteristics, and third by objects, that have not been in in the focus of attention yet.

The combination of these methods of attention shift enables the model to create complex search strategies through the production rules.

Results and Discussion

We present a simulation environment and a cognitive model of driver attention during car driving that is able to interact during run-time. In this work, two driver models were developed. The first model is able to reliably keep the traffic lane on a two-lane road and initiate a lane change followed by overtaking another vehicle. It identifies another vehicle and decides to overtake it if the safety distance falls below a certain distance (Fig. 6, scenario 1 and 2). The second model builds up on the first model and extends its functionality by identifying crossroad (Fig. 6, scenario 3), traffic signs and vehicle on the right hand side which have right of way (scenario 4, 5 and 6).

To obtain an integrated driver model of human driving behavior, it is essential to develop models in an architecture which is not task specific and can also model human behavior also in a different context, like ACT-R. This model is a first attempt to recognize, still simplified, traffic signs and crossroads. The development of an integrated driver model makes a first step towards the vision of accident-free driving. A majority (over 80%) of the automobile accidents are caused by the driver themselves. Fig. 1 shows the human errors while driving. Nearly 16% of the accidents happen while turning and during exit, followed by disregarding the right of way (15%) and not-adapted speed (15%). Theoretically, the cognitive driver model could give a deeper insight for around 30% of the human errors while driving. However, it has to be taken into account that the model is still interacting with a simplified environment and not yet taking into account driver´s prior experience, which could be implemented by an increased attention in potentially high accident risk situations. Our driver model is one approach to integrate operational (lower-level) and tactical (higher-level) aspects in the framework of the ACT-R architecture. The model and the environment do not present a complete picture of driver behavior yet, but they form a base to extend the ETA framework in any direction.

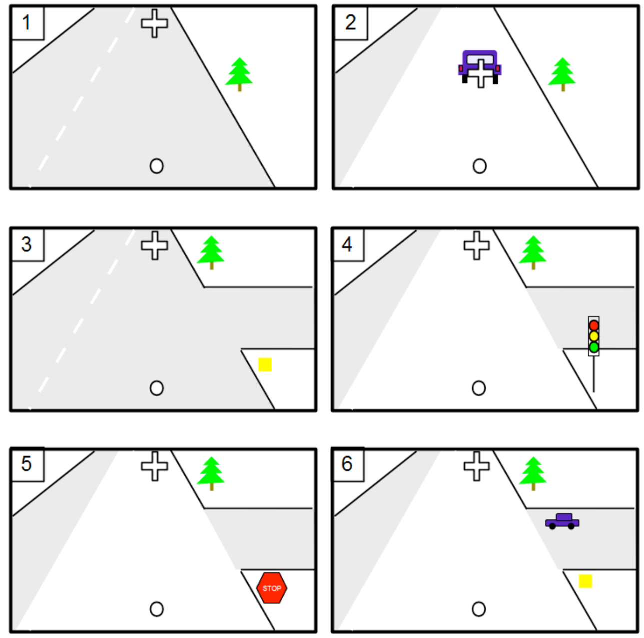

Fig. 6: Standard situations while driving a car which can be handled by the cognitive drive model.

The aspect of limited cognitive resources is one of the main factors for the adequacy of the model. Based on the implemented bottleneck, the three components control, monitoring, and decision making, have to share cognition. If the model is occupied with attention shift, it cannot simultaneously update the near point. Also, the model can only fire on production rule at a time and only one visual operation can be executed at a time. These processes take a certain time. For example, in the standard implementation in ACT-R, one firing of a production rule requires 50ms. This enables the researcher to compare the produced data with human data, because the ACT-R architecture produces an output file. This file contains the time, the active buffer and the according event. This study did not validate the model data so far. Future research could compare the output file data with human data, specially compare the attention shift of the model to human drivers over eye-tracking and the reaction times. However, for this validation, it must be possible from the technical side to either connect the ACT-R model directly to the simulation environment or to produce the same output file for the human data as the model does. Also, only most critical parts of key scenarios can be validated as no single method is sufficient enough to understand the complex task of human driving behavior yet.

Conclusion and Outlook

The progress to date in the development of cognitive architectures has been impressive, yet scientific gaps, technical challenges and practical issues remain. On one hand, cognitive models help to develop an understanding of driver behavior and aim to provide a theoretical account for human attention while driving. On the other hand, they are powerful and practical tools when implementing human-centered design and real-world applications. First steps towards the examination of the source of human mistakes through distraction from the primary driving task through secondary tasks like dialing a phone haven been taken (Salvucci, 2001) showing the feasibility of the architecture for these task and possible extensions.

The ACT-R architecture enables to elucidate interesting aspects and provides a theory of human attention while driving. At the same time, human attention during driving is a challenging task for the ACT-R cognitive architecture. It shows the still existing limitations beyond basic laboratory tasks and pushes the research community to expand the architecture towards more complex and finally real-world tasks.

Acknowledgments

We are grateful for the kind advice and assistance of Prof. Bernhard Nebel, the support of Prof. Wolfram Burgard (University of Freiburg), and for intensive discussions with Dario Salvucci (University of Philadelphia). This work has been partially supported by a grant from the DFG to MR (Project R8-CSPACE in the SFB/TR8 "Spatial Cognition").

References

Aasman, J. (1995). Modeling driver behavior in Soar. In: Leidschendam, The Netherlands: KPN Research

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C. & Qin, Y. (2004). An Integrated Theory of the Mind. Psychological Review, 111, 1036

Anderson, J. R. & Lebiere, C. (1998). The atomic components of thought. Mahwah, NJ: Lawrence Erlbaum.

Bellet, T., Bailly, B, Mayenobe, P., & Georgeon, O. (2007). Modelling Driver Behavior in Automotive Environments. Critical Issues in Driver Interactions with Intelligent Transportation Systems. Cognitive Modelling and Computational Simulation of Driver Mental Activities. Benmimoun, A. (2004). Der Fahrer als Vorbild für Fahrerassistenzsysteme? Ein fahrermodellbasierter Ansatz zur Entwicklung von situationsadaptiven FAS. 13. Aachener Kolloquium Fahrzeug- and Motorentechnik

Boer, E. R. (1996). Tangent point oriented curve negotiation. IEEE Proceedings of the Intelligent Vehicles 96 Symposium

Boer, E. R. (1999). Car following from the driver`s perspective. Transportation Research – Part F, 2, 201-206

Byrne, M. D. (2001). ACT-R/PM and menu selection: Applying a cognitive architecture to HCI. International Journal of Human-Computer Studies, 55, 41-84

Donges, E. (1987). A two-level model of driver steering behavior. Human Factors, 20, 691-707

Fajen, B. R., & Warren, W. H. (2003). Behavioral dynamics of steering, obstacle avoidance, and route selection. Journal of Experimental Psychology: Human Perception and Performance, 29, 343-362

Gray, W. D., & Altmann, E. M. (2001). Cognitive modeling and human-computer interaction. International encyclopedia of ergonomics and human factors, 1, 387-391, Taylor & Francis, Ltd.

Land, M., & Horwood, J.(1995). Which part of the road guide steering? Nature, 3, 77, 339-340

Liu, Y. (1996). Queuing network modeling of elementary mental processes. Psychological Review, 103, 116-136

Liu, Y., Feyen, R., & Tsimhoni, O. (2006). Queuing Network-Model Human Processor (QN-MHP): A computational Architecture for Multitasking Performance in Human-Machine Systems. ACM Transactions on Computer-Human Interaction 13, 37§70

Llewellyn, L. (1971). Visual guidance of locomotion. Journal of Experimental Psychology , 91, 245-254

Michon, J. A. (1985). A critical view of driver behavior models: What do we know, what should we do? Human behavior and traffic safety, 485–52, Plenum Press

Pomerlau, D., & Jochem T. (1996). Rapidly adapting machine vision for automated vehicle steering. IEEE Expert, 112, 19-27

Reid, L. D., Solowka, E. N.,& Billing, A. M. (1981). A systematic study of driver behavior steering control models. Ergonomics, 24 , 447-462

Rushton, S. K., Harris, J. M., Lloyd, M.R., & Wann J.P. (1998). Guidance of locomotion on foot uses perceived target location rather than optic flow. Current Biology, 8, 1191-1194

Salvucci, D. D. (2001). Predicting the Effects of In-Car Interface Use on Driver Performance: An Integrated Model Approach. International Journal of Human-Computer Studies, 55, 85-107

Salvucci, D. D. (2006). Modeling Driver Behavior in a Cognitive Architecture. Human Factors, 48, 362-380

Salvucci, D. D., Liu, A. , & Boer, E. R. (2001). Control and monitoring during lane changes. Vision in Vehicles, 9

Salvucci, D. D., & Gray, R. (2004). A Two-Point Visual Control Model of Steering. Perception, 33, 1233

Statistisches Bundesamt (online 20.12.2011) www.destatis.de

Wilkie, R. M., & Wann, J. P. (2003). Controlling steering and judging heading: retinal flow, visual direction and extra-retinal information. Journal of Experimental Psychology: Human Perception and Performance, 29, 363-378