A Car Driver’s Cognition Model
D. Krajzewicz, R. Kühne, P. Wagner,
Institute of Transportation Research, German Aerospace Centre, Berlin, Germany
Daniel.Krajzewicz@dlr.de, Reinhart.Kühne@dlr.de, Peter.Wagner@dlr.de
Tel.: +49 (0)30 67055204, Fax. +49 (0)30 67055202

1. Abstract
There is a basic need in transportation planning and traffic engineering for developing and testing traffic models of different granularity. Although our major interest is the replication of traffic within larger areas, both the current research on traffic safety and the desire to improve the quality of microscopic simulations makes it necessary to deal with the car driver’s cognition on a finer scale. This paper presents our model assumptions for such sub-microscopic simulations, which are based on results from cognitive psychology.

Although some preliminary work of this type is available, most of these applications are not open to the public, which makes them useless for scientific purposes. The cognition simulations available up to now mostly deal with memory processes and are not easily extendable by further structures such as vehicles with their dynamics or a representation of the simulated environment. These considerations motivated us to develop the above mentioned model from scratch.

The design of the model described herein includes sub-models of a human being’s perception, visual attention, internal environment representation and decision making as well as the execution of actions in a simulated vehicle. Results both from cognitive psychology and the research on human-machine interaction are incorporated.

This paper reveals our premises for a driver’s cognition model and describes the model itself, followed by a discussion of the model’s restrictions. As the implementation process is not yet closed, only some basic results are presented and a look into the future of the model is given.

2. Introduction
Due to road traffic’s high complexity, simulations are a major tool for traffic assessment and forecasting. While used for some decades already, traffic flow models have evolved considerably in the last 15 years. These changes put the driver more into the focus. In early research, traffic was mostly simulated using macroscopic flow models, which simulated traffic in analogy to gas or fluid. Then, the growing computation speed available allowed the usage of microscopic models even for larger areas. Although already known before, microscopic models of traffic flow became popular within traffic simulations, especially due to their ability to model the situation on intersections – details which macroscopic simulations were not able to model at all. As microscopic simulations mostly model traffic flow by calculating a set of car-driver instances, the driver’s behaviour gained bigger interest.

Further areas that make considering the driver in detail are the development of new assistance systems and the research on traffic safety. Both fields use the driver as the focus. While the research on the physics of traffic - the origin of most microscopic traffic models - looks from outside at a car-driver-object’s dynamics, these topics try to consider the processes within the driver. They mainly regard the restrictions of a human being’s mind, in order to identify and solve the problems arising from these restrictions.

Having some experience with microscopic simulations and due to the importance of the topics safety and assistance systems, we were interested in what a driver is really doing when driving. A further goal was to develop and validate new microscopic models at the finest level by evolving larger, sub-microscopic models first. These sub-microscopic models shall incorporate detailed models of the driver and the vehicle. After their validation, we want do derive the most important influences of the process of driving.

At first, we tried to find a model based on the research on a driver’s cognition that fulfilled the following requirements: it should be described mathematically and implemented as a computer program; it should be available as open source in order to make its extension and exploration possible; and it should be easily extendable, both by new models of the cognition’s sub parts and by different vehicle models. We found several candidates that may be divided into two classes: 1. Full-featured man-in-the-loop simulations including models for the human operator and the vehicle he steers, such as MIDAS [1] or PELOPS [2]. 2. Algorithmic models of a human’s cognition itself, such as COGENT [3] or ACT-R [4]. In fact, none of the candidates was appropriate for our purposes: The applications belonging to the first class, although very complete and elaborate, are used within the developing companies only, not allowing to take a look into the application’s interior and to validate the model. The programs or libraries belonging to the second class are mostly reduced to model basic memory operations. Attributes located at the boundaries of human cognition, such as visual perception and the execution of actions are not regarded herein, or if,
in a too simplified way. It is also quite difficult to incorporate vehicle dynamics into these models. The only model that we thought to be applicable, ACT-R, was not open for extensions at the time our project started. We assumed that the extension of ACT-R by a vehicle model and an environment model would be more time extensive than implementing an own application. In the meantime, a new version of ACT-R became available for free and is already used for traffic purposes by Salvucci and colleagues [4]. Before presenting the model itself, we will discuss possible fields of research that sub-microscopic models may be used for within the next section. The description of our model will then follow. This report presents some in-work experiences, as the development is ongoing. A discussion of missing structures will follow. We will discuss this issue after the model has been described and we will conclude with a short overview of the results available so far.

3. Capabilities and Limitations of Sub-Microscopic Traffic Models

While macro- and microscopic simulations are used to predict traffic within the considered network, one cannot state this for sub-microscopic simulations. Their execution time is still too high to use them for large simulation areas where the movement of several hundred vehicles shall be investigated. Some possible fields for their usage may be instead:

- Investigations of the used vehicle’s ergonomics
  - Space available within the vehicle
  - Visibility of the environment a vehicle offers
  - Ability of the driver to understand the presentation of information within the vehicle
- Investigations of the road network’s ergonomics
  - Visibility of parts of the road network
  - Ability of the driver to understand the presentation of information along roads
- Investigations of traffic accidents
  - Development of quantitative models
  - Explanation for certain accident types
  - Forensic accident analysis
- Investigations of inter- and intra-personal driver variability
  - Driver disabilities
  - Influence of alcohol and drugs
  - Influence of emotions
- Explanation for macroscopic traffic states
- Population of driving simulators with other vehicles
- Validation and presentation of new systems (in-vehicle and on the road)
- Convincement of decision makers

Although these desires justify the development of a sub-microscopic model, it is clearly not possible to investigate all of them using the same framework. We will now discuss some typical problems.

Some of the mentioned research topics need very elaborate models that require a long time to be executed. An example for this is seen in attempts to validate the visibility of a vehicle’s displays. A model for such a computation was given by Delacour et al in [5]. Due to the large amount of computations to be done, a simulation incorporating it would run slowly. On the contrary, if someone wishes to use a sub-microscopic model in order to populate a driving simulator with traffic, real-time simulation is necessary. We decided to make our simulation run in real time. The neglect of modelling a vehicle’s display visibility was one thing we have withdrawn for this reason.

The execution time is, of course, not the only scale that is of importance for the model’s design. On the one hand, the model is meant to a) incorporate as many cognitive processes that are needed to resemble the driving process as possible and b) to simulate the driving process as close to reality as possible. On the other hand, it should be easy to understand and extendable by new mechanisms and structures.

Some of the processes that take place within the driver’s mind are not yet known in a way that makes a simple usage of the research results applicable for computer models. To model them, one has to use own approaches instead. To reduce the model’s complexity, some parts of a human being’s information processing were not modelled. We will describe them within the presentation of the model itself. These reductions were done because of the simulation speed and the complexity of the needed models.

Much investigation has been done to model the cognition in a consolidating way. Most traffic simulations use two control models, one for following of the vehicle in front and one for the lateral control of actions such as lane
changing. Such a cognition split cannot be found within the cognitive literature; it is rather assumed that all actions are controlled by a single instance. Our model of the cognitive processor will be described in the next section.

4. Model Overview

The task to comprehensively simulate information processing requires a proper modelling of the environment that surrounds the driver. The scene description used is made out of single objects which all have a three dimensional geometry. Further attributes for the appearance are used, but only for visualisation purposes. This allows not only the visualisation of the scene, but also the retrieval of visual information in a way similar to reality.

Often, the process of driving is described as a control loop. Jürgensohn ([6]) gives an in-depth description of this paradigm and we want to describe it only briefly here. When driving, the driver perceives the environment through his sensors. Only the eyes, the ears and the perceived acceleration forces seem to be relevant for driving. Due to some limitations of the sensors, only a part of the information available at the entry of the sensors is made available to the cognition’s processor. Also, before being passed to this processor, information is filtered by an attention instance. Both the sensors’ limitations and the fact that information is being filtered are joined within the next picture to the “sensor-filters” structure.

The processor uses the gained information together with general knowledge about the world and knowledge about the route to derive the actions that shall be executed next. It steers its sensors by setting the attention to those parts of the environment that are most important to continue the process in a secure way. The derived actions are passed as commands to the muscular system. Due to some circumstances, such as haste, imperfections of the body control, etc. these commands are distorted. This is visualised in the picture as “action filters”. The actions that the driver performs affect the controls of the vehicle he uses. The loop is closed by the vehicle, which moves in accordance to the settings of its controls, changing its position within the environment.

Due to the model size, we can not describe every detail herein. Instead of trying to describe all the influences from cognitive psychology that have been taken into account, we will concentrate on some certain sub-models. We hope to present some interesting solutions for some of the problems raised by sub-microscopic models currently under development (see [7] and [8] for examples).

4.1. Sensors

4.1.1. Visual Perception and Visual Attention

The functioning of the visual apparatus is quite well understood, at least for the first steps of the information processing done herein (see [9] and [10]). The physical structures of human eyes are well investigated as well as some of the first layers of the brain which process the seen light. Still, when coming to deeper regions, especially those that assign meaning to the seen objects, no valid models exist, precluding a simple reimplementation of the brain functions. A further problem for modelling of perception is the fact, that the information processing done within the brain is performed in a massively parallel way. This makes it impossible to implement these processes on a normal, sequential computer if fast execution is wished. Due to both limitations, we had to implement an own model of the perception. Instead of using the paradigm of information extraction from the environment by an
increasing decomposition of the seen scene into single objects, we use the one mainly found within simulations: the assignment of information to the objects by their explicit declaration within the simulation settings. Within our simulation, each object knows whether it is a vehicle, a street, a building etc. This information, together with a certain object’s attributes, is passed to the cognitive processor and both the reception times and a human beings’ limitations are taken into account by simulating them as time- or value-dependent filters.

During a single time step the position on the retina is computed for each object within the scene. If the object is visible for the simulated driver, this object’s distance from the centre of the retina is computed in order to get the quality it is seen with. This quality decreases from the middle of the eye, the fovea, to the retina’s boundaries. The simulation knows which object’s perception is computed and is able to determine this object’s identifier and type. As a human being needs some time to recognize an object or one of its attributes, the values, being fully computable within the simulation, are filtered before being passed to the visual attention instance and the cognitive processor. These filters take into account the position of the object on the retina and the duration of the object’s visibility. In contrary to processing of visual information, a human being’s cognitive processor works sequentially. The concentration on the perception of a single thing, be it an object, a set of objects, or an attribute of a single object at a time, is known as visual attention and much research has been done in this area (see [11] for an example). Most of the research on visual attention is made by measuring the eye movements (see [12]). The duality of the attentional processes, being either filtering uninteresting information and/or explicitly concentrating on a desired one, was incorporated into the model. A single object is taken into account by the cognitive processor and started to be considered. The information about other objects is suppressed and gets into the processor in a more distorted way.

4.1.2. Audio Perception

Most of the information needed for leading a vehicle is gained through the eyes. Still, it is known that the driver uses his ears to estimate his vehicle’s speed and to predict the time he has to switch the gears. Our model of the auditory sensors is kept very simple. The vehicle’s sub models for the engine and the tires incorporate values for the sound they produce, which may be retrieved by the simulated driver. No other influences are taken into account at this point.
4.1.3. Acceleration Perception

It is known that each driver has his preferred acceleration forces, both when choosing his longitudinal (see [13]) and lateral (see [14]) speeds. As the vehicles’ power increased over years, these driver preferences constrain the speeds of vehicles more than the vehicles’ own abilities.

Within our model, acceleration forces are not used as values the driver receives, but as values constraining his desired speed and acceleration in a certain situation. We will discuss this in more detail in the description of the cognitive processor that uses these constraints. The lack of a real perception of the acceleration forces disables the model to simulate what happens if the desired forces are exceeded. We will discuss this issue in a later section.

4.2. Central Executive

4.2.1. Information Processing

The human mind is a complex network of neurons and each of them is both an information processor as it gives information further to other neurons and an information storage by having a certain electrical potential. A clear distinction between both data and data processing is not given at the physiological level. This makes a proper translation into the data and algorithms – the paradigm used in computer programs – difficult. The usage of artificial neural networks was not a solution for us, as they are very slow in execution and furthermore do not allow making the processes understandable. Nevertheless, besides the modelling of the data a human brain uses, one has to model their processing.

Our model of the central executive contains the following sub-parts described in detail within the next subsections. These are: 1. An internal world representation – a model of a human being’s knowledge about the current state of the world that surrounds him. 2. The plan – a hierarchical structure where the complete route and current actions are stored and maintained. 3. Schedulers for actions and for the spread of visual attention. As actions known to the driver are stored in a certain structure, too, one could say that we also use a model of the long term memory, but the processes of information retrieval from the LTM are far too unknown to use this nomenclature seriously.

4.2.2. Internal World Representation

The human mind does not only work on information that has been gained from the world, but always also uses knowledge about the perceived objects and their behaviour it has learned before. The compound information about an object stays available for a short time in order process, even if the object is no longer in sight. The storage for this information is known as the internal world representation.

Within the sub model of the internal world representation, each currently interesting object is stored together with its current attributes. Furthermore, the simulated driver’s assumptions about this object’s further states are stored if the object is a dynamic one. These assumptions are used to determine where collisions may occur.

![Schematic view of placeholders (archetypes) for lanes (left) and vehicles (right) that describe a situation within the internal world representation](image)

The internal world representation is implemented as a set of placeholders, one for each current and next lane in route, the previous lane, the neighbour lanes, and the lanes that cross the driver’s route divided by whether they are foes to the driver or not. A similar distinction is done for the vehicles. A human being does not operate on continuous directions, but rather on direction prototypes (see [15]), and our model clearly discriminates between eight directions a second vehicle may be located at. A similar approach was already proposed for traffic simulations in [16].
Additionally, we use a further placeholder that may contain more than one vehicle to describe vehicles crossing the driver’s road. If one of the “informational archetypes” does not exist in the current environment as, for example, when the road the driver uses has only one lane, its placeholder stays empty. If filled, vehicle-placeholders contain the vehicles’ current and anticipated dynamics. Lane-placeholders indicate the knowledge about the lane’s shape the driver has to follow. This view on the internal representation is based on the following assumptions. When driving, each object has its own meaning or role. A driver reacts to the vehicle in front of him in a different way than he would if the vehicle was behind him. Vehicles on neighbouring lanes are relevant mainly if the driver decides to change the lane. They are not as important as the vehicle in front. The vehicle in front determines the vehicle’s speed strongly, as known from vehicle following models. The reason why we model crossing vehicles from both directions in the same way is that their role stays the same independent of the direction they come from: either the simulated driver has to wait or not, in dependence on the rules of the approached junction. Using this approach, we avoid holding all objects around the driver unsorted, which would need more effort to evaluate the situation.

4.2.3. Plans
The cognitive processor works on the current values within the internal world representation. In order to perform actions that allow him to continue his route, he also has to keep his desired route in mind. While driving, the simulated driver decides about lane changing, his desired velocity, and the amount of turning of the steering wheel in order to follow curves. These actions are performed by the simulated driver in dependence on the situation on the current road. On a further level, lane keeping and acceleration/deceleration processes are performed.

Our model resembles the three levels of vehicle control described by Michon and revealed by Ranney in [17]. The navigation layer is used as a knowledge base about the desired route. At the start of the simulation, this route information is given as a list of points to pass. These points are given as abstract positions naming the subset of the junctions to pass, sorted by their occurrence within the route. Information about the desired direction at each junction is given to the action-planning instance that calculates the actions needed, as soon as the junction is visible. On the last level, control actions such as lane keeping are done. These actions are not implemented as discrete events as those located on the operational/control level, but as a set of controlling equations and are not planned, but executed in respect to the current stimuli. They will not be described within this report.

We will describe the tactical layer more deeply, now. As every action contains premises, such as a maximum speed possible to perform it, the instantiation of an action such as turning on a junction implies a set of further required actions. The next image shows how actions are initiated by a simple desire to turn left on a junction. The vehicle approaches the junction on the right lane. The desire to turn is set on position (1) as the left lane is the only one he is allowed to use. As the driver has to choose a velocity that allows him to turn safely, on position (1) he has not only to begin turning, but also have a proper velocity. This indicates that he has to break at a certain distance, position (2) herein. He also has to change the lane in order to get to the left one. Normally, this is done at a secure distance. This action is visualised as (3) within the image. A lane can only be changed if the driver is sure not to violate other traffic participants. The needed perception of the environment is done at position (4). At least at position (5) the opposite
traffic must be regarded. At position (6) – after leaving the curve – the driver may accelerate to the speed allowed on
the following lane.

![Visualization of actions and their premises using turning left as an example](image)

The origin of this model is the assumption that each action is stored within the human mind as a template to be
parameterised before being executed, called “schema”. The current research on this topic, stemming mainly from
cognitive psychology, assumes a schema to hold the following information (translated and slightly modified from
[18]): a) the motor program for this action, b) the initial situation the action may be started within, c) the result
reached by executing the action and d) the input from sensors during execution of the action. We assume also the
knowledge about an action’s duration to be of importance. This knowledge is implicitly stored within the part which
describes an action’s results. Knowing the premises and the result of an action, the driver can choose the action that
should be performed within the current situation. Furthermore, before executing an action, the driver can predict the
action’s duration and dynamics and incorporate them into a plan.

On a more abstract level, each of these positions seems to describe things to be done at a certain position within the
driver’s path and the driver’s assumptions about dynamics at a certain position. We assume that describing the
position using a spatial position only is not sufficient. Rather, some actions seem to be executed if a certain event
occurs. Due to this, we assume these tactical control points to be a projection of the following form: (event|distance)
→ (state|action initialisation). One may notice that we use “distance” rather than “areal position”. By doing this, we
hope these points to be valid to describe a car driver’s car-following behaviour, which is rather determined by the
distance to the leading car rather than to certain positions in space. Due to their characteristics, we name these
tactical points “event distance hooks” (EDHs).

This model of action execution appears to be quite powerful. Setting of action execution points to describe actions to
perform in the future not only resembles a driver’s possibility to anticipate the situation, but also to plan his actions.
We do not have enough experience with the model, yet, to describe it more precisely. Some questions remain,
especially those that concern action scheduling. They are described within the next subsection.

4.2.4. Action and Attention Scheduler

When driving, the driver sometimes gets into situations that require him to change his plans on a finer time scale. If
the driver wanted to turn right at the junction he approaches, but the vehicle in front of him stops, he has to overtake
this vehicle first. All actions the driver plans to perform have a certain importance within the current situation.
Michon’s model of three levels of vehicle control tries to give an explanation about the dependencies between
actions located at different levels, but does not consider how actions on the tactical level are scheduled. Within our
model, a certain structure is responsible for action scheduling.

This action scheduler is currently in development. The following issues seem to be of importance.

a) What happens if an action cannot be fulfilled? b) How can scheduling of actions be designed in order to replicate
both an action’s importance and urgency? c) How is it possible to generalize the invocation of actions from all
objects to be considered? The current implementation allows setting actions needed within the next time steps, but is
not yet flexible enough to allow their easy removal or update.

Information about perceived objects stays within the internal world representation for a certain time. If not updated
by taking the object in sight again, they are forgotten and get lost. The driver has to update periodically the
information he needs to drive safely. If the situation gets more complicated – several objects have to be evaluated –
the driver has to plan his sensory actions. If free, the attention is directed towards the next interesting object. We
assume “interest” to be a function of both the importance and the urgency with which an object has to be taken into
sight. Furthermore, on the scheduling level, the knowledge about the duration of evaluating the object and the
knowledge about the time the object will stay known is taken into account by a human mind’s action planning
instance.
4.3. Action Execution

The simulated driver acts by moving his hands and feet to a desired position that is retrieved from the action he wants to perform. If, for example, the action “push throttle” to 40% is executed, the right foot is moved onto the throttle, if not already being there, and to the position that sets the throttle to be pressed by 40%. No further parts of the body are simulated, as a too complicated body model would be required for this. Some sub-microscopic man-in-the-loop simulations, MIDAS ([1]) for example, do use body models. We hope our approach to be exact in a desirable way as it is capable to compute whether one of the extremities is already involved in an action and to simulate the times needed to bring the affecting extremity to the desired position. These values are needed for simulating of action scheduling processes and for simulating of the movement of body parts over time, respectively.

5. Discussion on the Limitations

Within the model description, we have mentioned some neglected parts that a complete human cognition model could have. Due to these limitations, our model is not applicable to some of the topics listed within the section about general applications for sub-microscopic traffic models. The neglect of the haptic feedback precludes the model from simulating the acceptance for driver assistance systems that use haptic feedback to report certain vehicle states to the driver. Such methods are used within active cruise control or heading control systems (see [19] and [20]), where the driver gets information about these systems’s ideas about a correct speed or steering wheel angle via haptic senses. The neglect of these topics is in accordance with our wish to concentrate specifically on the dynamics of the driving process. This is also true for the neglect of the body as one can assume the vehicle’s interior is appropriate for a human driver. In accordance to this is a missing sub model for the perception of the vehicle’s displays. Possibly, this can not be stated for neglecting of the perception of acceleration forces, as the driver may wish to change his speed when the forces get dangerous. Although the model assumes correct wishes about speeds in curves and their proper resembling by the simulated driver, the simulation of a less experienced driver who is not capable to adapt the desired forces appropriately may become inadequate and an extension of the model would be needed.

6. Results and Further Work

The model is still under development and some validation has to be done before we can present any tangible results. Still, some things that raised attention during the implementation should be mentioned. Most microscopic traffic simulations do not regard the driver’s favoured acceleration forces. This is something that not only influences the dynamics of when the vehicle starts – on traffic lights for example, but also during the passing of tight curves. Furthermore, the car-following paradigm does not fit to situations within the real world. Beside the leading vehicle, a driver has to take into account other objects or attributes, such as the next junction’s right of way rules, and has to do it in an anticipatory way, which means that these do influence the driving process before they are reached.

On the cognitive scale, this model shows how the scheduling of driving by a single cognitive instance can look like. We are convinced that the concept of the EDHs is appropriate for this and should be investigated further in the future. Furthermore, we assume the structures described herein to include all processes needed to model a vehicle driver completely.
7. Conclusion
Our aim has been to present a cognitive car driver model that incorporates most of the known issues and that should be applicable to most modern problems. Beyond incorporating some cognitive structures that have not been addressed within the traffic research to date, such as the internal world representation or a complete plan instance, its special feature is the usage of a single, hierarchical model for both the longitudinal and lateral movement.
As the verification is not closed yet, we neither state its validity nor present any further implications on the traffic research than we made within the results section. Our goal is to complete and validate the model by the middle of 2004.

8. Acknowledgements
We would like to thank Dr. Thomas Jürgensohn from Human Factors Consulting for his advice on our work as well as other people this work is influenced by. Furthermore, we want to thank Markus Hartinger and Collin Vance for reading this work and their corrections.

9. References
[13] “Car-following: a historical review”; M. Brackstone, M. McDonald; in: Transport Research, 1999; Elsevier; (pp. 181-196)