Toward Adaptive Support: Modeling Drivers' Allocation of Attention

R. Kovordányi, S. Ohlsson

Department of Computer and Information Science
Linköpings Universitet
SE-581 83, Linköping, SWEDEN
E-mail: ritko@ida.liu.se

Abstract
Driver distraction and inattention are major contributing factors in traffic accidents (cf., e.g., Najm et al, 1995). Some of these accidents might be avoided in the future if drivers' (miss)allocation of attention could be detected, and the driver be prompted toward key events in the traffic scene. Our long-term objective is to develop a cognitively based driver model where drivers' allocation of attention can be simulated for diagnostic purposes. We present a connectionist model framework implemented in PDP++, embodying the notion that focusing attention on a visual object is closely coupled with an intention to use or act upon this object, for example to steer away from it. We intend to continue our work with exploratory simulations centring on the human motor system. Our present objective is study to what extent multiple tasks can be parallelized by drivers. For further information, visit http://www.ida.liu.se/~ritko.html.

Keywords
Distraction, Inattention, Cognitive driver model, Adaptive driver support

1 INTRODUCTION

In-vehicle devices, such as built-in phones and navigation aids can easily distract drivers by capturing their attention and diverting their gaze from the traffic scene. In other situations, drivers can become inattentive by inappropriately choosing to be preoccupied with phone dialing, instead of paying full attention to driving. When a driver is momentarily incapable of detecting driving-related visual events because he/she is preoccupied with non-driving-related tasks, the driver can be said to be *temporarily* overloaded. One way to avoid overload is to adapt device-driver-interaction to the driver's momentary capacity to handle non-driving-related tasks. Systems for adaptation of in-vehicle device messages are becoming standard in certain car models. However, present systems are not capable of taking the drivers' *cognitive* load into account—adaptation is instead based on external variables, such as variations in traffic intensity—nor can they support the driver when he/she becomes distracted by non-device-related events, or is simply inattentive.

In this article we propagate for an approach to adaptation based on monitoring, through real-time simulation, of the driver's cognitive processes that is, processing of visual information, and cognitive management of multiple simultaneous tasks. We use the term ‘task’ to refer to *the cognitive processes* underlying operational-level driver actions (Michon et al, 1990), such as looking over the shoulder, or turning the steering wheel. The question we would like to answer is: For any moment in time, will the driver be able to allocate his/her *attention* to an unexpected visual event, for example, a pedestrian that suddenly appears on the side of the road?

2 THE COGNITIVE APPROACH

As a first step toward this goal, we have implemented a cognitively based neural network architecture in which simulated attention can be allocated to individual visual events based on the driver’s current tasks. By employing a biologically-based model architecture, simulated allocation of attention can be constrained in the same way as in the human cognitive system (cf., e.g., Vidulich and Wickens, 1985). The point is that a driver model built on this architecture will exhibit a restricted ability to handle multiple tasks. This in turn allows for a realistic simulation of the driver’s cognitive workload.

Somewhat simplified, the human cognitive system can be subdivided into three mutually interacting subsystems: The perceptual, the central cognitive, and the motor system. While perceptual information is processed in parallel, motor movements can only be directed toward one object at a time. This creates a basic incompatibility: The parallel inflow of perceptual information must at some point be sequenced in order to provide an unambiguous basis for motor planing (Kovordányi, 2002). Hence, the driver will only be able to consciously perceive a limited amount of information at a time. The pieces of information he/she will become aware of will be influenced by the tasks that he/she is engaged in.

3 EXISTING DRIVER MONITORING SYSTEMS

Present approaches to driver monitoring and modeling are mainly based on statistical methods for situation assessment and/or statistical classification and prediction of tactical-level driver behavior (Michon et al, 1990), such as passing, turning, starting, and stopping. For example, MIT’s smartCar project adopts coupled hidden Markov models. Using this technique, the driver’s behavior can be recognized as constituting the initial action within a particular tactical-level maneuver category, such as lane change, or right turn. By recognizing the initiated maneuver, the driver’s action can be extrapolated into a prediction of the driver’s next actions (Oliver and Pentland, 2000).

While the smartCar project adopts hidden Markov models, Onken and Feraric (1997) present a driver monitoring system where a two-layer ART-network is used to learn and categorize the driver’s normal behavior. This system is able to detect deviations from the normal pattern, with the rationale being that such deviations might signal potential driver overload.

Both of these systems incorporate a driver model and implement driver behavior categorization by applying purely statistical methods to externally observable behavior. In contrast to this, Salvucci and Macuga (2001) have developed a *cognitive* driver model, which simulates how the driver handles two simultaneous tasks, namely steering within a lane and cell-phone dialing. A simplifying assumption made in this model is however that drivers’ allocation of attention corresponds to overt (externally observable) eye movements.

4 THEORETICAL BASIS FOR THE DRIVER MODEL

We take a stance in the notion that attention is a *cognitive* resource that has no simple correspondence to overt behavior. We subscribe to a late-selection-for-action view on selective attention (see Kovordányi, 1999, for a theoretical discussion and review of empirical evidence). In accordance with this view, Ballard and coworkers (1997) and Bugmann (2001) propose that allocation of visual attention might play a key role in motor movement planning. These and other researchers (Fagg and Arbib, 1998) suggest that attending to the visual form of an object provides the motor system with the necessary parameters that are needed for preparing the appropriate movement toward this object—for example, gripping an object with a 10 cm diameter requires a corresponding hand aperture.

4.1 Detection of visual events

The conscious registration of visual events can be assumed to require allocation of attention. This notion is supported by experimental data on 'inattentional blindness', referring to the fact that people often fail to notice events they are not prepared to see. Simons (2000), for example, describes an experiment where subjects were shown a video footage of a basketball game and asked to count the number of passes made by one of the teams. The experimenters let a man dressed as a gorilla pass by in this scene. Although the gorilla was visible for 5 seconds, 50% of the subjects did not notice it! Hence, preoccupation with one task (counting the number of passes) may have prevented subjects from becoming consciously aware of an unexpected event.

As attention tends to be allocated to the part of the visual space that contains the object of an intended movement (Fagg and Arbib, 1998; Bugmann 2001), it is conceivable that a driver will not allocate his/her attention to a particular visual event, unless he/she has formed an appropriate expectation or initiated response preparation for this event. However, we have to note that the evidence summarized above mainly concerns the coupling between the perceptual and motor systems as implemented via the "where/how" pathway (cf. figure 1). Allocation of attention along the "what" pathway (object-based attention) remains to be explored in our model. In addition, salient events, for example, movement in the peripheral visual field, are known to cause 'stimulus-based capture' of selective attention.

4.2 Allocation of attention vs. eye movements

While the choice of appropriate motor movements seems to require *allocation of attention* to the object of this movement, the final execution and fine-tuning of the same motor movement often relies on foveal input, and thus requires *eye movements* (Ballard et al, 1997; Bugmann, 2001). This entails that when a driver is fixating the target for one motor action, his/her attention may already be focused on another object, which is the target for the next action. Hence, at least for a certain kind of visually surveyed actions, a driver's eye movements, once they are executed and therefore are observable, are bad predictors for where the driver focuses his/her attention.

Given the problem of time-lag between observed movements and focus of attention, and hence the low predictor-value of observable actions, we intend to approach driver monitoring by employing real-time simulations of a cognitive driver model. As evidenced in our previous discussion, we consider the coupling between allocation of attention and motor planning, and hence the detailed functioning of the motor system, to be central to this approach.

5 THE IMPLEMENTED MODEL FRAMEWORK

In the following sections, we discuss the architecture and function of the motor system within the envelope of our model framework (figure 1). This framework is designed to simulate drivers' coping with two simultaneous tasks: steering within a lane and reacting to an unexpected visual event. These tasks are presently represented in our model by a continuous tracking task, and a discrete reaction task (Martin-Emerson and Wickens, 1992; Kieras and Meyer, 1997). The model is implemented as a limited parallel processing architecture written in LEABRA++ (O'Reilly and Munakata, 2000), a biologically realistic artificial neural network tool package within PDP++.

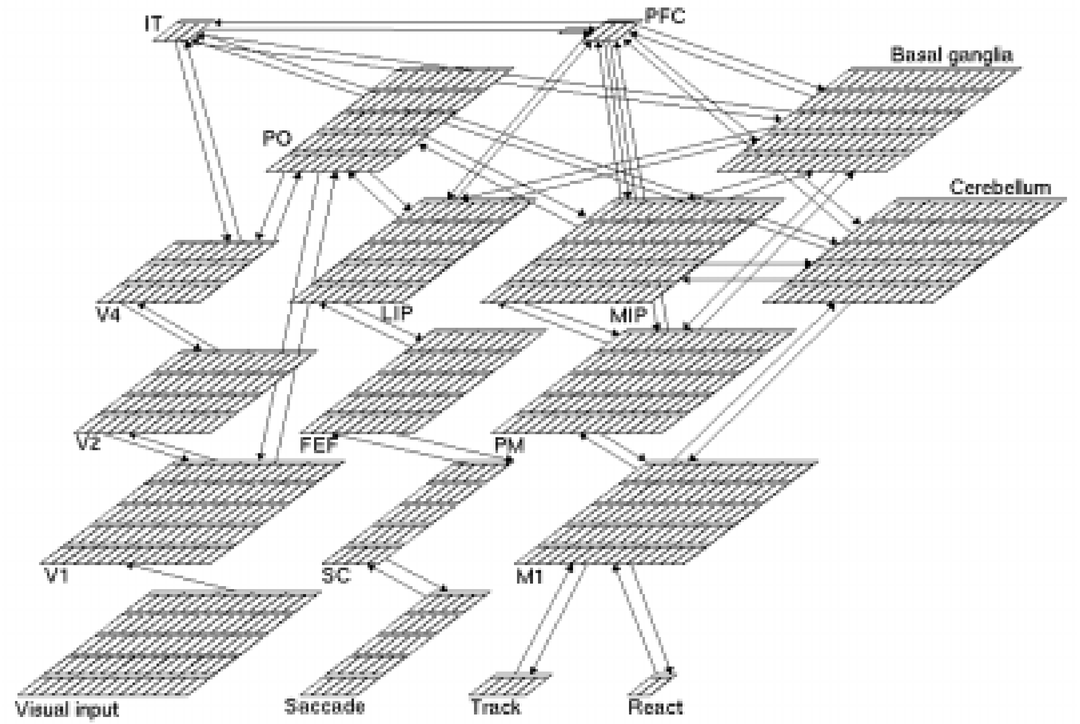

Figure 1. Present architecture of the driver model. V1-V4 denote visual areas in the "what" pathway, PO: parieto-occipital cortex ("where" pathway), PFC: prefrontal cortex, MIP and LIP: medial and lateral intraparietal cortex, FEF: frontal eye fields, SC: superior colliculus, PM: premotor areas, M1: primary motor cortex, BG: basal ganglia. Basal ganglia is assumed to play a role in the sequencing and timing of motor plans, while the cerebellum might be involved in choosing between alternative plans on the basis of visual input.

5.1 Motor planning

Our objective is to explore the coupling between allocation of attention and motor planning in order to find out to what extent multiple tasks can be parallelized. Of particular interest is the question of which preparatory phases in motor planning can be overlapped with another task—it is, for example, conceivable that the execution of one task can go on in parallel with the early planning phases of another task. Below, we present a number of hypotheses on the role of the cerebellum and the basal ganglia in the sequencing and timing of motor actions, with particular focus on the mechanisms for suspending and resuming a task (Kieras and Meyer, 1997).

Motor actions can be divided into three phases. The first phase is the selection of a motor plan. The next phase specifies the parameters for motor actions and the last phase is the execution of action. Salvucci (2001) also recognizes two distinct parts of execution in the motor system where the preparation of a movement can be cancelled if cognition redirects and calls for another preparation of action. The second part can however not be cancelled.

Cerebellum and basal ganglia are assumed to be involved in both the control of attention and in motor planning. The challenge for research is on *how* these sub-cortical structures contribute to shifts of attention and executive control operations (Ravizza and Ivry, 2001). Bugmann (2001) implies that the cerebellum may have a role in preparing the sensory motor system for producing a task specific response. Posterior parietal cortex (MIP and LIP in our model) is supposed to initialize movement plans and cerebellum chose among these plans on the basis of sensory information. Hence cerebellum is thought to have a modulator role where it selects already learnt stimuli-response rules rather than generating and projecting new mappings (Bugmann, 2001).

The basal ganglia may play a role in the timing and sequencing of movement phases, as well as higher-level motor actions. Basal ganglia has a good regulating position since it projects inputs to the inferior temporal cortex, motor cortex and prefrontal cortex. Basal ganglia may also implement gating of perceptual input during motor activity (Rosenbaum, 1991). This suggests that new sensory driven motor preparation can not be initiated unless the current motor plan is suspended. One way this may be implemented has been suggested by Berns and Sejnowski (1996). They propose that the timing difference between the two main pathways of the basal ganglia block competing motor programs via another indirect pathway.

In summary, while the cerebellum is involved in the coupling between the perceptual and the motor system by choosing motor plans on the basis of sensory information, the basal ganglia may play a role in implementing executive control operations.

6 FUTURE WORK

As a direct continuation of the work reported here, we plan to explore alternative hypothesis on how motor planning, in particular how the suspend/resume functions may be realized in the basal ganglia. We intend to validate the resulting model using empirical observations on dual-task performance (e.g., Martin-Emerson and Wickens, 1992), involving a continuous tracking and a discrete reaction task.

The modeling methodology we intend to use is exploratory: As opposed to simply demonstrating the adequacy of a specific model, we intend to find out which model components are crucial for obtaining a good fit to empirical data, thus providing a more reliable basis for subsequent real-time simulations (Kovordányi, 2000; 2001).

References

Ballard, D. H., Hayhoe, M. M., Pook, P. K, and Rao, R. P. N. (1997) Deictic codes for the embodiment of cognition. *Behavioral & Brain Sciences*, 20(4), 723-767.

Berns, G. S., and Sejnowski, T. J. (1996): How basal ganglia make decisions. In A. R. Damasio, H. Damasio, and Y. Christen (Eds): *The neurobiology of decision making*. Berlin: Springer Verlag, 101-113.

Bugmann, G. (2001): Role of the cerebellum in time-critical goal oriented behavior: Anatomical basis and control principle. In S. Wermter, J. Austin, and D. Willshaw (Eds): *Emergent neural architectures based on neuroscience.* Lecture Notes in Computer Science. London: Springer Verlag.

Fagg, A. H., and Arbib, M. A. (1998): Modeling parietal-premotor interactions in primate control of grasping, *Neural Networks,* 11(7/8), 1277-1303.

Kieras, D. E., and Meyer, D. E. (1997): An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human-Computer Interaction,* 12, 391-438.

Kovordányi, R. (2002): Sequencing of information versus interfacing between processing levels. *Computational Intelligence,* 18(1), 47-49.

Kovordányi, R. (2001): Factorial modeling: A method for enhancing the explanatory and predictive power of cognitive models. In E. M. Altmann, A. Cleermans, C. D. Schunn, and W. D. Gray (Eds): *Proceedings of the 2001 Fourth International Conference on Cognitive Modeling.* Mahwah, NJ: Lawrence Erlbaum Associates.

Kovordányi, R. (2000): Controlled exploration of alternative mechanisms in cognitive modeling. In *Proceedings of the Twenty-Second Annual Meeting of the Cognitive Science Society,* Philadelphia, PA.

Kovordányi, R. (1999): Modeling and simulating inhibitory mechanisms in mental image reinterpretation? Towards cooperative human-computer creativity. *Linköping Studies in Science and Technology,* Dissertation no. 589, Linköping University.

Martin-Emerson, R. and Wickens, C. D. (1992): The vertical visual field and implications for the head-up display. In *Proceedings of the 36th Annual Symposium of the Human Factors Society*, 1408-1412.

Michon, J. A., Smiley, A., and Aasman, J. (1990): Errors and Driver Support Systems. *Ergonomics,* 33(10/11), 1215-1229.

Najm, W. G., Mironer, M. S., and Fraser, L. C . (1995): Analysis of target crashes and ITS/countermeasure actions. *ITS AMERICA Fifth Annual Meeting,* Washington, DC. http://www.itsdocs.fhwa.dot.gov/\JPODOCS\REPTS_TE/2W@01!.PDF

Oliver, N., and Pentland, A. P. (2000): A Driver Behavior Recognition and Prediction in a SmartCar. In *Proceedings of AeroSense 2000, Enhanced and Synthetic Vision, Orlando.*

O'Reilly, R. C., and Munakata, Y. (2000): *Computational explorations in cognitive neuroscience: Understanding the mind by simulating the brain.* Cambridge, MA: MIT Press.

Ravizza, S. M., and Ivry, R. B. (2001): Comparison of the basal ganglia and cerebellum in shifting attention. *Journal of Cognitive Neuroscience,* 13(3), 285-297.

Rosenbaum, D. H, (1991): *Human Motor Control*, San Diego, CA: Academic Press.

Simons, D. J. (2000). Attentional capture and inattentional blindness. *Trends in Cognitive Sciences,* 4(4), 147-155.

Vidulich, M. A., and Wickens, C. D. (1985): Stimulus-central processing-response compatibility: Guidelines for the optimal use of speech technology. *Behavior Research Methods, Instruments and Computers,* 17(2), 243-249.