# MODELING DRIVER BEHAVIOR WITH DIFFERENT DEGREES OF AUTOMATION:

A Hierarchical Decision Framework of Interacting Mental Models

# Ewin R. Boer<sup>1</sup> Marika Hoedemaeker<sup>2</sup>

<sup>1</sup> Nissan Cambridge Basic Research, 4 Cambridge Center, Cambridge, MA 02142, USA erwin@pathfinder.cbr.com

#### ABSTRACT

An integrated driver model is presented in which cognitive constructs such as driver needs are causally connected to the orchestration of skill based driving tasks. The proposed three layer hierarchical structure is composed of satisficing decision makers who communicate with intermediate lavers of dynamic mental models. The decision makers direct the information flow and decide which mental model is consulted and/or activated and when. Mental models on the other hand provide information at different levels of abstraction that guides the decision making process. With the aid of this hierarchical driver model, predictions can be made about how automation of particular driving subtask may influence the overall driving behavior and to what degree these may be driver dependent. We particularly focus on the effects of introducing an Adaptive Cruise Control (ACC) system by exploring the dynamics of the driver's ACC mental model as a function of experience with the ACC.

**Keywords:** Driver model, mental models, adaptive cruise control, satisficing decision making; human-machine interaction.

# INTRODUCTION

Driving can be characterized as goal directed behavior that is propelled by aspirational factors and obstructed by constraining factors. The adopted compromise between these opposing factors shapes the emergent behavior. The goal <sup>2</sup> Delft University of Technology, De Vries van Heystplantsoen 2, 2628 RZ Delft, The Netherlands Hoedemaeker@tsm.tudelft.nl

directed nature of driving is characterized by a set of higher level needs whose interaction affects the way in which drivers orchestrate the set of observable low level driving tasks. The proposed model structure is similar to the knowledge-, rule, and skill based control levels of Rasmussen [Rasmussen '83] and Michon's structure of three levels of driving tasks: strategical-, tactical-, and operational [Michon '85]. To characterize the means by which drivers derive at a decision in the face of conflicting needs we adopt satisficing decision theory [Goodrich, Stirling, & Frost '98]. To decide between alternative strategies, decisions, or actions, the decision maker is assumed to consult mental models which evaluate the consequences of committing to a particular alternative. To operationalize the notion of mental models, we adopt the following definition [Boer & Liu '97]:

A mental model is the internal representation employed to encode, predict, evaluate, and communicate the consequences of perceived and intended changes to the operator's current state within its dynamic environment.

In addition, these mental models are assumed to operate at different levels of abstraction [Moray '98].

With the aid of the proposed hierarchical driver model, predictions can be made about how automation of parts of the driving task may influence behavior and to what degree these may be driver dependent. We particularly focus on the effects of introducing an Adaptive Cruise Control (ACC) system by exploring the dynamics of the driver's ACC mental model at different levels of abstraction. When an ACC system is available and therefore restructures the car following task, a very important question is whether this has any potentially adverse safety effects. The primary concern is that drivers may not have or have not yet developed an adequate mental model to predict when or under what circumstances the ACC may not be able to perform the car following task adequately, for instance when hard braking is necessary.

When drivers drive with automated systems, they construct mental models of these systems which leads them to either make correct or incorrect predictions about the performance of the system. The beliefs and inferences about the system will be based upon the design of the system and their experience with it. They lead to context specific expectations about system behavior thus influencing the driver's interaction with the system (e.g. whether they should reclaim control or not). Therefore it is important that the driver develops the appropriate mental model of system functionality [Stanton & Young '98].

In order to insure that drivers develop the appropriate mental model of the systems they interact with, we believe that a fundamental understanding of the effect that automation has on the driving experience is very important thus motivating the development of an integrated driver model.

#### **DRIVER MODEL**

Human decision making and action taking are grounded in satisficing decision making rather than optimality seeking. In natural and often complex situations humans adopt those strategies that are adequate rather than optimal for reasons such as that not enough time is available to evaluate all alternatives<sup>1</sup>, or that only incomplete knowledge of the situation is available which renders comparison of all alternatives not justifiable, or that optimality can not be defined because the utility measures appropriate for the different alternatives are incompatible<sup>2</sup>. Humans choose from the strategies that come to mind or that present themselves through other means the one that is good enough. In satisficing decision theory good enough is defined as the set of decisions (actions) for which the total benefit is greater than the total cost. This notion of adequacy is expanded on below and is described in detail in [Boer, Hildreth, & Goodrich '98].



Figure 1 Hierarchical Driver Model. MM = Mental Model, DM = Decision Maker.

The hierarchical framework of driver behavior is composed of satisficing decision makers who communicate with intermediate mental models (Fig. 1). The decision makers direct the information flow and decide which mental model is consulted and/or activated and when. Mental models on the other hand provide information that facilitates decision making by guiding the perception, prediction, evaluation, action cycle of performing a particular task. At the tactical level, the decision maker takes the role of attention manager and task scheduler [Boer, Hildreth, & Goodrich '98]. The information flows through the proposed hierarchical structure as follows:

At the highest level, need satisfaction is evaluated. At that level, it is assessed whether the current trip choice is satisficing (i.e. good enough; see below for formal definition). Assessments of degree to which various needs are met is provided by the performance mental model. If the current route is not satisficing, the route mental model is consolidated to provide evaluations about the expected consequences of

<sup>&</sup>lt;sup>1</sup> As explained in the companion paper [Boer, Hildreth, & Goodrich '98], one of the advantages of satisficing decision making over optimal decision making is that alternatives can be evaluated independently to assess whether they are acceptable or "good enough". An immediate consequence is that an exhaustive search is generally not required because upon finding the first acceptable alternative (which is offen the one currently active), the search can be terminated thus promoting efficient use of limited attentional resources. This is not so in optimal decision making since all alternatives need to be compared in order to find the best one.

<sup>&</sup>lt;sup>2</sup> In satisficing decision theory each alternative is evaluated independent of other alternatives. This offers the flexibility to considder alternatives that are evaluated based on widely different and possibly incompatible criteria. Such an evaluation may not be feasible in optimal decision making (except in a contrived manner) because all alternatives need to be converted to a common performance measure.

taking alternate routes. The usefulness of these evaluations depends on knowledge about and past experiences with those routes.

At the strategic level, a decision is made about which route (perhaps the current one) is good enough. This choice is communicated to the route mental model which predicts upcoming situations and feeds them to the tactical decision maker for anticipatory attention management and task scheduling.

Given the current situation and the set of mandatory and discretionary driving tasks, the task scheduler at the tactical level divides attention and resources over the different low level operational tasks handlers. Each operational task has its own mental model, that continuously perceives the state of nature (state of driver, vehicle, automation, and environment), evaluates the consequences of alternative decision (including the one currently active), predicts future states of nature in response to various alternatives, and communicates this information to the decision maker who uses it to manage attention and schedule relevant tasks [Boer, Hildreth, & Goodrich '98].

Which tasks or maneuvers are considered is also influenced by the strategic level. If a particular route does not yield satisficing performance, but an alternative route is not available, then the task set (Fig. 1) may be expanded with tasks that may, for example, improve expediency such as lane changes and overtaking maneuvers. If the trip is satisficing without these tasks, then the driver does not need to evaluate situations for windows of opportunity to perform these tasks thereby enhancing efficient use of attentional resources.

The proposed structure implies that there is a relationship between drivers' needs and their actual driving styles. Driving styles reflect habitual modes of operating the car on the road [West & French '93]. Hoedemaeker [Hoedemaeker '96] investigated this relationship between needs and driving styles. The driving styles were based on the Driving Style Questionnaire (DSQ) by West et al. [West el al. '92], and the relation between needs and driving styles was investigated for three kinds of trips: a routine trip of commuter traffic, a long distance trip, and a short distance trip. The results show a relationship between needs and some of the self reported driving styles. They also show that needs are a stable trait that can differ between drivers, but not between trips, and that driving styles do differ within a driver for different trips.

# Strategic Level: Needs and Strategy Selection -

In the proposed hierarchical decision structure, the strategic level represents how a driver's particular needs for a given trip are used to determine whether the driver is satisfied with the way the trip progresses. Needs are higher level goals that drivers try to satisfy (below we characterize them by a set of criteria relevant to the driving experience). According to Rumar [Rumar '93], the motives and the goals for traveling are the basis for the needs. The primary goal for drivers is of course to reach the destination, but they do not accept that goal at any price. They aim to achieve their goals to a high degree of satisfaction without violating constraints too much. The driver requires a certain time, speed, safety, economy and comfort. These are needs that drivers try to satisfy. This is not always possible because of the current traffic situation. Needs are therefore translated into local goals that are congruent with the situation at hand. So it is not the needs that are adapted but the local goals that are adapted to the current situation. For example, if speed or driving as fast as possible are a driver's most important need, then this can be translated into driving fast with many risky lane changes, taking a different route with less traffic that allows for higher but illegal speeds, or taking a commuter lane even though the driver is alone in the car.

In order to cast the effect of needs into a computational framework, the following taxonomy is proposed:

- Risk or Safety
- Expediency (speed, driving time, reaching destination as quickly as possible)
- Pleasure (favoring certain routes or roads, enjoying the surroundings)
- Kick of driving
- Workload
- Economic cost
- Compliance to social norm

We realize that this taxonomy has some complicated inter-dependencies but feel that they are so highly situation dependent that it is still useful to treat them independently and let the driving context dictate how they are correlated and how these correlations affect decision making. Moreover, truly independent criteria in complex naturalistic decision making are very difficult to come by<sup>3</sup>.

<sup>&</sup>lt;sup>3</sup> One could resort to principle component analysis to

Some of the needs can be seen as global goals to achieve or as benefits, whereas others function more as constraining factors or as costs. In driving, benefit (accuracy<sup>4</sup>) is composed of *expediency (EX), pleasure (PL),* and the *kick of driving (KD)* which characterize a driver's global goals or reasons for getting in the car in the first place. The cost (rejectability) are those criteria that constrain behavior and include *risk (RK), workload (WL), compliance to social norm (SN),* and *economic cost (EC).* They constrain the degree to which these goals can be achieved.

Formally, the satisficing set of decisions  $D^{S}$ is the one for which accuracy A(D) exceeds rejectability  $R(D)^{5}$ . In words, the set of decisions that bring one closer to the goal without incurring too much cost are called satisficing. The satisficing principle can be applied at all three levels in the driver model hierarchy. In this paper, we primarily focus on the strategic level. In our companion paper [Boer, Hildreth, & Goodrich '98] we adopt satisficing decision theory as the basis for decision making at the tactical decision level where attention management and task scheduling take place. In that paper, we also outline how decision making or action taking at the operational level is influenced by changes in the importance of or weights assigned to each of these seven criteria at the strategic level.

To determine at the strategic level whether a particular strategy (including the current one) is satisficing, the composite accuracy and rejectability utilities are computed. For demonstration purpose, we adopt a linear weighting of accuracy (rejectability) criteria to derive at the composite accuracy (rejectability) utility. To assess the particular means by which the various criteria are combined, further experimentation with human subjects and computer simulation of a computational implementation of the driver model is required. Thus, the accuracy associated with a particular strategy  $D_i$  is obtained with<sup>6</sup>

$$A(D_i) = w_{EX}^A EX(D_i) + w_{PL}^A PL(D_i) + w_{KD}^A KD(D_i)$$

where  $EX(D_i)$  for example, is the numerical representations of the expediency utility that a driver attributes to strategy  $D_i$ . Similarly, the composite rejectability utility value attributed to decision or strategy  $D_i$  is<sup>7</sup>

$$R(D_{i}) = w_{RK}^{R}RK(D_{i}) + w_{WL}^{R}WL(D_{i}) + w_{SN}^{R}SN(D_{i}) + w_{EC}^{R}EC(D_{i})$$

Thus, the satisficing set of decisions is defined as the set for which accuracy exceeds rejectability

$$D^{S} = \left\{ D_{i} \middle| A(D_{i}) > R(D_{i}) \right\}.$$

In case this set is not singleton, one still needs to select one from the set of satisficing alternatives. This is accomplished by applying what is referred to as a tie-breaker. Several different tiebreakers can be formulated which each yield slightly different final results. These details are not relevant to this paper but the interested reader is referred to our companion paper [Boer, Hildreth & Goodrich '98].

If the driver is not satisfied (i.e. the current strategy is not a member of  $D^s$ ), strategic decisions are made in terms of changing routes or considering an expanded set of tasks such as lane changes and overtaking which may otherwise not be considered at every instance especially if the current trip is satisficing thus avoiding the need to repeatedly expend attentional resources to look for windows of opportunity in which these maneuvers may be performed. Mental models intermediate to the strategic and tactical level provide estimates of the expected utility of these alternatives task sets or routes to assess whether they may yield satisficing performance. Note that performance evaluation at the strategic level takes place on longer time scales than those at the tactical or task management level because they require, as

establish independent variables but the interpretation of these variables is expected to become highly situation dependent thus defeating the reason for establishing principle components in the first place.

<sup>&</sup>lt;sup>4</sup> Accuracy and rejectability are often used in the satisficing formalism. Accuracy criteria (utilities) in the driving context can be interpreted as those characterizing reasons for getting in the car in the first place or the goal one tries to acomplish during the trip. Rejectability criteria characterize the constraining forces that limit the degree to which these goals can be satisfied. In satisficing decision theory, the tradeof between these two opposing sets of criteria results in a set of alternatives (on time scales ranging from decision about whether to overtake or not to decisions about whether to take one particular route or not).

<sup>&</sup>lt;sup>5</sup> We adopt one particular theory of satisficing decision theory of which a formal account is given in [Goodrich, Stirling, & Frost '98].

 $<sup>^{6}</sup>$  We use the symbol D to denote decision. Whether the decision is made between different strategies, tasks, or actions depends on the level at which the decision is made in the driver model hierarchy. Here we focus on the strategic level.

 $<sup>\</sup>overline{}$  Without loss of generality it is assumed that the individual utility functions (e.g. on expediency or risk) are normalized and that the weights not only reflect their relative weighting but also the mapping to a common reference frame so that the different terms can be combined.

explained below, aggregation of multiple performance measure at various time scales to derive at a stable estimate of the utility of the various criteria.

Once a particular route has been selected, the route mental model guides the task management level, via a situation mental model, by providing turn by turn instructions (e.g. turn left at next intersection). The task manager updates, through the situation mental model, the route mental model each time a particular maneuver has been completed.

#### Tactical Level: Attention Management and Task Scheduling

At the tactical or rule based level the *task* manager orchestrates which maneuvers or low level driving tasks (e.g. lane keeping, car following) are performed (Fig. 2). Again, we adopt a satisficing approach to determine which skill based controllers should be invoked at what times to assure an acceptable performance level [Boer, Hildreth, & Goodrich '98].



Figure 2. Detailed structure of the communication that takes place from and to the tactical task management level.

The task scheduler communicates performance related information, via а performance mental model, to the strategic level. The performance mental model consolidates performance measures related to the driver's seven global needs, to derive at a numerical representation for each of them (see below for one possible computational implementation). The following list provides some insight into the types of measures that influence the value of the

seven needs (criteria) that shape the driving experience<sup>8</sup>:

- *Risk:* e.g. how frequently was it necessary to switch to a critical event task handler (e.g. obstacle avoidance, lane correction, hard braking); average time headway; total time over which time-to-collision was less than 4s;
- *Expediency*: e.g. average speed; distance traveled;
- *Pleasure*: e.g. amount of time attention was directed to the environment or in-vehicle devices such as a CD player;
- *Kick of Driving*: e.g. percentage of time in which full acceleration, high speed curve negotiation, or fast driving was possible;
- Workload: e.g. fraction of time that attention had to be allocated; frequency with which tasks were switched;
- Economic Cost: e.g. how much time was spent in stop and go traffic, city driving, mountains, etc. (compute flue usage); average speed; distance traveled; how many tickets;
- Conformance to Social Norm: e.g. how often other drivers beeped their horn; frequency with which a lane change was made without leaving a sufficient time headway for the trailing car in the target lane;

Many of these terms  $T_j$  (e.g. average speed, frequency of task switching) are running statistics (with forgetting) which are compiled in the performance mental model (Fig. 2). Formally, for a given high level need or criterion  $C_i, k \in [1,7]$ , its value is computed using

$$C_k = \sum_{j=1}^{N_k} c_{kj} T_j$$

where  $N_k$  is the number of terms  $T_j$  that contribute to criterion  $C_k$ , and  $c_{kj}$  is the weight of the *j*-th term in the *k*-th criterion<sup>9</sup>. Note that

<sup>&</sup>lt;sup>8</sup> This list of terms that contribute to the utility value assigned to a particular criterium is by no means complete and requires careful experimentation to establish which is topic of current research at Nissan CBR. In [Levison & Cramer 95] and account of similar terms is provided for their penalty minimizing (i.e. based on optimal decision making) driver model.

<sup>&</sup>lt;sup>9</sup> Again, the linear combination rule is applied for demonstration purposes only. The "true" combination rule can only be obtained through careful comparison between results obtained with human subjects and an implementation of the driver model. We specifically do not mention the use of high focussed lab experiments under highly constrained conditions because the strategies that humans adopt under those conditions do generally not teach us much about the

the same term may contribute to multiple criteria. Without loss of generality, the weights are such that  $0.0 \le C_k \le 1.0$ . An alternative formulation is that  $C_k$  is simply the running average of the criterion value of the decision selected from the satisficing set at the task managers level<sup>10</sup>. An advantage of the former is that it maintains the flexibility that different aspects of performance and experience play a different role at different levels. This is very clear in case of the low level controllers where notions of workload and economic cost, for example, are often difficult to define. The ultimate goal is to derive at the most natural representation of subjective performance (e.g. what may be considered satisficing) at each level in the driver model. Performance is used here in general terms as a measure of goal accomplishment whereby goals are clearly very different at different levels.

# Operational Level: Low Level Task Handlers and Skill-Based Controllers

Every single basic operational driving task (e.g. car following, lane keeping, over taking, intersection negotiation) has its own mental model that guides the perception-action cycle. Monitoring, integration, and evaluation of these subtasks scheduled by the task manager [Boer, Hildreth, & Goodrich '98]. The set of active mental models is dependent on the situation. Depending on the situation different mental models ask for different amounts (frequency and duration) of attentional resources. Switches between mental models occur when: i) the

strategies driver's adopt in naturalistic setting. One reason that optimal control and optimal decision models has been relatively sucessful is because they were often applied to situations in which the human subjects were highly motivated, very well trained, had a very clearly defined goal, and a limited set of choices compared to the real world. In naturalistic decision making, these behavior constraining conditions fall away and a much broader spectrum of behaviors emerges. This is one of the reasons why we are actively exploring the use of satisficing decision theory rather than optimal decision theory because satisficing decision theory is capable of modeling the natural variability in people's choices because of its set based nature which explains why people adopt very different strategies or exhibit very different behaviors, or make very different choices, or apply very different control under apparently identical conditions because to them all these alternatives are good enough. <sup>10</sup> This alternative formulation in which each low level task

<sup>10</sup> This alternative formulation in which each low level task performance and execution is directly characterized in terms of these seven criteria is elaborated on in [Boer, Hildreth, & Goodrich '98]. situation demands a new maneuver to be executed or a new task to be performed, ii) the current maneuver, which has been left in open loop or unattended mode for some time, can no longer be guaranteed safe thus calling for a perceptual and possible control update, iii) spare attentional resources are available and some task that does not require steering or pedal input can be initiated.

Given that driving can be considered a process of carefully orchestrating an array of mandatory and discretionary tasks, the issue of appropriate attention management and task scheduling is of critical importance. This issue in particular plays a central role in our driver model (Fig. 2). Low level mental models are assumed to communicate their need for attentional resources to the higher level attention management system. The task manager requests information from the low level task handlers or skill based controllers and assigns attentional resources based on who needs it most. It performs time and resource multiplexing between the low level task handlers or skill based controllers. Each operational mental model is assumed to communicate the following information to the task manager:

- Performance evaluation (to be communicated to higher levels);
- How many resources are required to perform the task;
- At what time in the future will be resources be needed and for how long.

In our companion paper [Boer, Hildreth, & Goodrich '98], we discuss in detail those aspects of attention management and task scheduling that are relevant to the orchestration of low level task handlers as well as the ways in which top down information from the strategic level affects these delicate interactions.

#### INTERACTION WITH ADAPTIVE CRUISE CONTROL

The introduction of automation does not significantly alter the above described dynamics, it merely introduces an extra mental model associated with the task that the automation is capable of supporting or performing. Several interesting issues arise with the introduction of a system that is capable of vehicle control under a limited set of conditions<sup>11</sup>. Insertion of an

 $<sup>^{11}</sup>$  At this stage of vehicle automation, most systems fall in the category of driver assist systems. They do not claim to

automation system offers the task manager an additional choice, namely whether to use the automation or do it manually. Three examples of different types of vehicle automations are

- 1. Assist driver in performing a particular low level task (e.g. power steering, antilock braking).
- 2. Automatic attention management (e.g. information and warning systems)
- Automatic performance of a particular low level task (e.g. automatic car following or lane keeping)

Adaptive Cruise Control (ACC) systems can be seen as an enhancement of the conventional cruise control systems and automate part of the car following task. The systems use radar or laser range finders to track vehicles in the forward field and automatically maintain a preset following distance to the vehicle directly ahead traveling in the same direction. The system is understood to be primarily a convenience feature for the driver instead of being able to avoid collisions when the car in front is making an emergency stop or a vehicle cuts in front of the ACC vehicle with a very small headway (i.e. the deceleration range of the ACC is limited).

A mental model of ACC is initialized via instructions about its interface, its functionality, and its operational limitations which are subsequently generalized or restricted depending on individual differences and preconceived beliefs about automation [Parasuraman & Riley '97]. Interaction, exposure, and experience then shape the mental model of ACC through a feedback mechanism (Fig. 3) in which prediction errors are used to update the various components of the mental model (operational constraints, set of situations it can handle, etc).

During the period following the insertion of a device or system (ACC), a new mental model of the system is constructed and developed. Some of the important issues in this adaptation process are the level of automation or the degree to which the driver (operator) is taken out of the continuous control loop, the new role of the driver, the rate at which semi-critical events (the ones that the ACC cannot handle) develop, and the frequency with which these events tend to occur [Hancock, Parasuraman, & Byrne '96]. The driver needs to, ideally, take all these issues into account to arrive at the appropriate safety conscious role division between human and machine.



# Figure 3 Detailed structure of ACC mental model.

If the ACC is capable of performing the car following task safely and comfortably under all circumstances, then the driver can effectively disable the manual car following mental model. However, given the current technological limitations, ACCs will not be able to perform without human monitoring and intervention. It is important to consider how drivers may obtain and develop a mental model of the ACC and how they may use it. Ideally, the driver's mental model of the ACC should offer the task manager a realistic assessment of ACC performance as well as an estimate of how long it can be left unmonitored. This requires knowledge of:

- 1. Interpretation of the display interface,
- 2. Timeliness and reliability of warnings,
- 3. The range of situations and conditions the ACC can handle (speeds, time headways, time to collisions, weather),
- 4. An estimate of the dynamics of the traffic situation (provided by the driver's traffic mental model).

With a correct mental model of the automation, drivers can effectively assess whether manual control is favored over automation in a particular context. Ecological interface design is especially important in supporting automation since it explicitly shows the constraint boundaries of the

take over any part of the driving task for 100% or under all conditions. They merely assist the driver in performing a particular task with less effort. In nearly all cases, drivers are still expected to pay attention as when performing the (semi-)automated task themselves. For example, adaptive cruise control (ACC) systems, that regulate speed in order to maintain a particular time headway to the preceding vehicle, often have limited sensing capabilities and are often not equipped with emergency braking, thus requiring drivers to identify these situations are take precautionary action.

system's operational domain [Vincente & Rasmussen '92] thereby to some degree externalizing one of the roles a human operator's mental model of the automation plays. For these reasons, it is believed that ecological interfaces improve human-machine interaction and expedite learning.

Adaptation of the ACC mental model is the result of two processes that operate at different abstraction levels of the mental model. The first is based on the degree to which the expected behavior of the ACC in a particular context, as provided by the ACC mental model, differs from that observed. Depending on the degree of misprediction, the ACC mental model is updated to arrive at a more and more accurate account of the ACC's operational domain.

The second process that affects mental model adaptation takes place at a higher abstraction level in the mental model<sup>12</sup>. If it appears that adaptation of the ACC mental model does not reach a stable configuration because of apparent inconsistencies in the ACC's behavior (i.e. not easily tied to contextual differences), then the levels of trust, usefulness, and efficiency attributed to the system may decrease (Fig. 3). The result is that drivers may start to rely less on the system and more on manual performance of the driving task. If, on the other hand, the ACC mental model does converge and if the operational constraints are easily tied to a particular situation and conditions, then the driver may start to rely more on the system. Consequently, prediction errors may not be evaluated as frequently to allow for more efficient use of attentional resources.

Trust in automation is of paramount importance in the design of automation because lack of trust may result in abandonment of the automation [Muir & Moray '96]. Trust is a subjective measure of the degree to which the system can handle a set of situations reliably, safely, dependable, and effectively. Trust in our framework is a measure associated with the mental model of automation (ACC). It is assumed that trust is defined at a higher level of abstraction in the ACC mental model. The mental model of the ACC is primarily shaped through interaction with the physical system while trust is shaped through evaluations of the adequacy of the operational mental model of the ACC (i.e. no predictability without a model). The measure of trust can be regarded as an efficient means to assess whether ACC should be used in a particular situation (e.g. highly inhomogeneous traffic flow) or under particular circumstances (e.g. rain). This assessment can be supported by going down towards lower levels of the abstraction hierarchy, but such an assessment is generally not required when quick decision making is needed. Similarly, usefulness and efficiency are subjective labels based on evaluation of mental model predictions. They are efficient measures used in the process of attention scheduling and task management and therefore influence whether using the automation makes the overall task satisficing.

Crucial to the development and evaluation of any model that characterizes human behavior in environments is a means complex to experimentally gain access to the acquisition, structure, content, and use of mental models. This presupposes that the notion of mental models provides a practical foundation for describing human-machine interaction. If we adopt the definition of mental models given in the introduction, several avenues for experimental exploration present themselves (as an example we focus on how one may go about "measuring" the state and development of a driver's mental model over a period following the installation of an ACC).

A mental model is a dynamic entity which requires initialization. This can be accomplished via verbal or written instruction and can be formulated at many different levels of abstraction and in many different domains [Moray '98]. Depending on the exact formulation, this may instill a particular level of trust in the system. Note that drivers are assumed to be highly capable of performing the task manually.

A mental model is dynamically shaped by experience. This means that the situations encountered, the circumstances experienced, and the system responses observed alter the operational domain (set of contexts) within which the ACC is expected to perform safely.

The mental model is attributed with several criteria (e.g. trust, usefulness, and efficiency) at the highest level of abstraction. Drivers are hypothesized to assign context dependent utility values to these criteria that fluctuate during initial exposure but ultimately converge. These

<sup>&</sup>lt;sup>12</sup> To some extend the various abstraction levels of a mental model can be regarded as meta levels of lower levels. They are the result of many to one mappings [Moray '98].

utility measures may guide decision making as explained below.

The rate at which the mental model develops over time is a reflection of the driver's learning process. Ideally this learning process should be very short and converge to a mental model that is a representative characterization of the ACC system's abilities and limitations.

In short, one has experimental control over the initial content of the mental model and the way it develops during a period of exposure to the system and interaction with the system. One can probe the operator's internal representation of the system's operational domain by monitoring driver's interaction with the system as well as monitoring their attentional allocation (using eye movement as a proxy for attention). One can assess the degree of trust, usefulness, and efficiency that the operator places on the system's functioning within its expected operational domain by comparing a driver's task scheduling activities under different task loads<sup>13</sup>. Experimental exploration of the structure, content, acquisition, and use of mental models along these lines is topic of current research at Nissan CBR.

Since trust can be regarded as a utility value that drivers place on a particular system's performance in a particular context it can be used in a satisficing framework to characterize when drivers are willing to assign responsibility to the system. Lee and Moray [Lee & Moray, '94] showed that plant operators use automation when their measure of trust exceeds their measure of self confidence. This high level evaluation of whether to use automation or not without comparing alternatives fits the satisficing framework very well. In a multi task environment, an operator's self confidence about being able to perform a particular task depends not only on his or her ability to perform that one task but also on how many other tasks need to be attended to and how many tasks await scheduling. Just as trust is a measure associated with a mental model of a particular system

(automation), self confidence may also be regarded as a measure of trust associated with the mental model of ones own ability to perform that particular task combined with the measure of available attentional and physical resources. The degree to which these measures can be used to model task scheduling is topic of current research.

### CONCLUSIONS

An hierarchical structure of information processing is proposed in which mental models at different levels communicate with satisficing decision makers to perform goal directed behavior under the influence of internally and externally imposed constraints<sup>14</sup>. The goal is to develop an integrated driver model that offers a framework in which the effect of partial automation of the driving task can be characterized for different types of drivers under circumstances. different The adopted characterization of the driver's interaction with the vehicle and environment offers means to generate experimentally testable predictions and hypotheses. The proposed model is an attempt to capture the flexibility and efficiency of human drivers that are attributed to: i) the ability to switch between and operate at different levels of abstraction depending on the familiarity with a particular situation, ii) the adoption of a satisficing rather than optimizing decision making process to efficiently use attentional resources, iii) the employment of task specific skill based controllers that require minimal attentional resources, and iv) the use of mental models to guide when particular information needs to be sampled to aid continuous operation. Topics of further research are the computational implementation and experimental validation of this integrated driver model.

# REFERENCES

[1] E.R. Boer, E.C. Hildreth, & M.A. Goodrich (1998). A driver model of attention management and task scheduling: Satisficing decision making with dynamic mental models, In. Proceedings of the XVIIth European Annual Conference on Human Decision making and Manual Control, Valenciennes, France, Dec. 14-16.

<sup>&</sup>lt;sup>13</sup> It is expected that drivers will rely more and more on automation of a particular task as their task load (defined by the number of task that also require attention for successful completion) increases. By assigning cash costs and benefits (rewards) to the failure and success of each task, one creates a common measure of utility which can be used to quantify the utility value placed on the criteria of trust, usefulness, and efficiency associated with performing a particular task automatically. This utility value can be compared against the utility value placed on self confidence of being able to satisfactorily perform that task. We are currently designing experiments that follow this approach.

<sup>&</sup>lt;sup>14</sup> A similar approach of hierachical information processing structures is used in the design of intelligent control systems [Gupta & Sinha 96].

- [2] E.R. Boer & A. Liu (Eds.) (1997). Cambridge Basic Research 1997 Annual Report, Nissan CBR, Cambridge, MA 02142.
- [3] M.A. Goodrich, W.C. Stirling, and R.L. Frost (1998). A theory of satisficing decisions and control. *IEEE Transactions on Systems, Man,* and Cybernetics OPart A; Systems and Humans, Vol 28, No. 6, pp. 763-779, Nov.
- [4] M.M. Gupta & N.K. Sinha (Eds.) (1996). Intelligent Control Systems: Theory and Applications. New York: IEEE Press.
- [5] P.A. Hancock, R. Parasuraman, & E.A. Byrne (1996). Driver centered issues in advanced automation for motor vehicles. In R. Parasuraman & M. Mouloua (Eds.), Automation and human performance: Theory and applications, pp. 337-364. Hillsdale, NJ: Erlbaum.
- [6] M. Hoedemaeker (1996). Drivers' needs and behaviour. In: Handbook of the International Conference on Traffic and Transport Psychology, May 1996, Spain. DL: V-1787-1996. pp. 147-153.
- [7] J.D. Lee & N. Moray (1994). Trust, self confidence and operators' adaptation to automation. *International Journal of Human-Computer Studies* Vol. 40, pp. 153-184.
- [8] W.H. Levison & N.L. Cramer (1995). Description of the Integrated Driver Model. BBN Report No. 7840. BBN Systems, 10 Moulton Street, Cambridge, MA 02138.
- [9] J.A. Michon (1985). "A critical view of driver behaviour models: what do we know, what should we do?", In: Evans, L. & Schwing, R.C., *Human behaviour and traffic safety*. Plenum Press, New York.
- [10] N. Moray (1998). Mental models in theory and practice. Attention and Performance XVII, D. Gopher (Ed.), The MIT Press, Cambridge MA.
- [11] B.M. Muir & N. Moray (1996). Trust in automation: Part II. Experimental studies of trust and human intervention in process control simulation. *Ergonomics*, 39, pp 429-460.
- [12] R. Parasuraman & V. Riley (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors*, Vol. 39, No. 2, p. 230-253.
- [13] J. Rasmussen (1983). Skills, Rules and Knowledge: signals, signs and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 13, pp. 257-266.
- [14] K. Rumar (1993), "Road User Needs", In: A.M. Parkes & S. Franzén, Driving Future Vehicles. Taylor & Francis, London.
- [15] N.A. Stanton & M.S. Young (1998). Vehicle automation and driving performance. *Ergonomics*, Vol. 41, No. 7, pp. 1014-1028.
- [16] K.J. Vincente & J. Rasmussen (1992). Ecological interface design: Theoretical foundations. *IEEE Transactions on Systems*, *Man, and Cybernetics*, Vol. 22, pp. 589-606.

- [17] R. West, J. Elander, & D. French (1992). Decision making, personality and driving style as correlates of individual accident risk. Contractor report 309. Transport Research Laboratory.
- [18] R. West & D. French (1993). Direct observation of driving, self reports of driver behaviour, and accident involvement. *Ergonomics*, Vol. 36, No. 5, pp. 557-567.