

A Model of Human Brake Initiation Behavior with Implications for ACC Design

Michael A. Goodrich
Computer Science Department
Brigham Young University
Provo, UT, USA

Erwin R. Boer
Nissan Cambridge Basic Research
Nissan Research and Development, Inc.
Cambridge, MA, USA

Hideaki Inoue
Nissan Research Center
Nissan Motor Co., Ltd.
Yokosuka, Japan

Abstract

Many automobile manufacturers have recently included or will soon include Adaptive Cruise Control (ACC) systems in some vehicles. The operational limit of ACC-generated (hard) braking is a critical factor determining the human driver's interaction with the automation. Using a satisficing-based multiple mental model perspective, we generate a characterization of the natural onset of human-generated braking. We hypothesize that effective human interaction with an ACC system, that is, an interaction resulting in safe and comfortable transition from ACC to human control via human intervention, is achieved when (a) the onset of automated braking matches that of a skilled human operator and (b) a human driver can easily detect and interpret the operational limits of ACC-generated braking. This hypothesis, which is supported by experimental evidence obtained from both driving simulator and test track studies, implies that effective ACC designs should perform braking by either (a) matching human behavior or (b) augmenting human ability to detect and interpret the operational limits of ACC-generated braking through a surrogate such as a warning.

1 Introduction

Adaptive Cruise Control (ACC) is perhaps the feature of advanced vehicle systems that has been studied most. The reason for this study is that the technology required for ACC implementation is feasible given the current state of the art, and the complicated human factors for safe and effective ACC use are being unraveled. With these advances in technology and human factors, many vehicle manufacturers have recently included or will soon include ACC systems in some automobiles. In this paper, we discuss our experimental and theoretical observations of what elements influence the safe implementation of ACC systems, and discuss the risks associated with the misapplication of these elements. Specifically, we advance the hypothesis that safe and effective ACC design requires that the operational limitations of ACC-generated braking must be *detectable* and *interpretable* by human drivers, and motivate this hypothesis using a multiple mental model framework substantiated by experimental results.

The motivating force behind automating certain aspects of driving is that automation can perform certain driving skills more efficiently than drivers, whereas the driver is more efficient at other skills. For automation to be safe, it is not only

necessary that the driver can implement their share of the skill set, but also necessary that the driver can intervene when the automated skills are inappropriate. One approach to designing ACC and other human-machine systems is to identify how human drivers perform the task and then emulate their behavior (subject to safety and technological constraints). Such an approach assumes that

1. skilled, attentive, and motivated humans do the task efficiently and safely,
2. the automation can be designed such that it emulates human behavior, and
3. the driver's use of automation is enhanced by similarities between the automation and human skill set, particularly when transitions are made between automated and human control.

Our approach to describing, predicting, and enhancing driver behavior is to identify the set of skills that automobile drivers use to safely manage speed and interact with traffic. To manage speed and interact with traffic, we suppose that drivers use a set of learned skills. Our approach then identifies how one skill-based behavior is switched to another and how perceptual cues trigger such switches. This approach produces a computational model that emulates driver behavior and, by associating ACC behaviors with a subset of natural driver skills, can be extended to predict how the driver switches between manual and automated behaviors by detecting and interpreting the operational limits of the automation. These predictions are then supported by experimental results and evidence gathered from relevant literature. These results complete the pilot work presented in [2] where we developed a computational method for describing and emulating driver behavior using a multiple mental model framework.

This paper is written from a perspective that assumes that ACC systems are primarily intended to increase driver comfort. Based on this assumed position, we discuss how these systems can be designed such that they are safe.

2 Skill Switching: A Decision Theoretical Characterization

Driving can be organized into a set of skilled activities that are applied when afforded by the environment [1]. In this context, the term *activity* means the driver's actions on the vehicle (e.g., an activity is pushing the brake pedal or turning the steering wheel). Formally, a *skill* can then be defined as

a learned sequence of driver activities. For example, speed management skills include following a lead vehicle, regulating speed about a desired value, and braking to avoid a collision with activities of pressing the brake or accelerator pedal. The driver must map environmental cues into selected driver activities; an efficient way to perform this mapping is to employ a pattern of activities specific for a particular task, and then implement this skill when appropriate. This approach uses a task-specific mental model to determine which skill is appropriate for the circumstances. In reference to Figure 1, there are a set of skills that are appropriate for various circumstances. For speed management and traffic interaction, these skills are collected in a skilled longitudinal control block. These skills evoke real-time brake and accelerator pedal activities in a learned pattern of behavior.

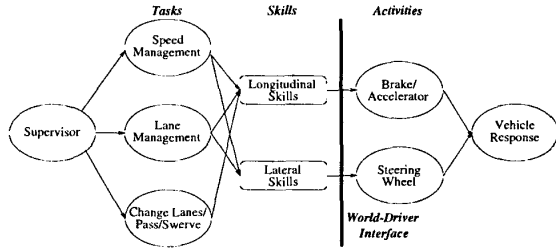


Figure 1: Skill switching via behavior-based reasoning.

For the speed management task, the set of skilled behaviors include free driving (regulating speed in the absence of other traffic), following another vehicle, and braking to avoid collisions. Thus, if U denotes the set of possible skills then, for the speed management task, the corresponding set includes $U = \{TR, SR, BA\}$, where TR indicates time headway regulation (car following), SR indicates speed regulation (free driving), and BA indicates active braking. Switches between skills are mandated when target perceptual states are not achievable by the currently enable skill-based behavior or when enabled skills are not acceptable for the given state. For example, a switch from speed regulation to braking is mandated when all lead vehicles stop.

For computational modeling purposes, skills can be emulated by closed-loop controllers that operate on environmental cues. These cues must be perceptually plausible meaning that drivers must be able to sense them. For the speed management task, vehicle speed v_A , time headway T_h , and time to collision (or, equivalently, inverse time to collision T_c^{-1}) are perceptually feasible perceptual cues (for a review of relevant literature, see [3]). We define the latter two cues, respectively, as $T_h = \frac{R}{v_A}$ and $T_c^{-1} = -\left(\frac{v_R}{R}\right)$, where v_A is the speed of the driver's vehicle, R is the range (relative distance between the driver's vehicle and a lead vehicle), and v_R is the rate of range change (relative velocity between the driver's vehicle and the lead vehicle). Given these perceptual values, a perceptual state can be defined as $\theta = [T_c^{-1}, T_h, v_A]^T$.

Satisficing Decision Theory (SDT) [4], which employs and compares two evaluation functions similar to the way bene-

fit and cost are compared in economics literature, is an ideal tool to describe switching between driver skills. In SDT, preferences over consequences are partitioned into a generalized type of benefit called *accuracy* meaning *conformity to a standard*, and a generalized type of cost called *liability* meaning *susceptibility or exposure to something undesirable*. Recall that U denotes the set of possible decisions or actions and Θ denotes the set of possible perceptual states. For each decision $u \in U$ and for each perceptual state $\theta \in \Theta$, a consequence results which is the effect of making decision u when nature is in state θ . The accuracy $\mu_A : U \times \Theta \mapsto \mathbb{R}$ and liability $\mu_L : U \times \Theta \mapsto \mathbb{R}$ set membership functions are preference relations defined for each consequence (i.e., action/state-of-nature pair). Any consequences which are more accurate than liable are acceptable, whence we can define the satisficing set as $S_b = \{(u; \theta) : \mu_A(u; \theta) \geq b\mu_L(u; \theta)\}$.

Given this SDT decision rule, we can restrict attention to those perceptual states which are satisficing for a given u , and those skills which are satisficing given the state of nature, respectively defined as $S_b(u) = \{\theta : \mu_A(u, \theta) \geq b\mu_L(u, \theta)\}$ and $S_b(\theta) = \{u : \mu_A(u, \theta) \geq b\mu_L(u, \theta)\}$. In terms of behavior management by a driver, suppose a skill $u \in U$ is being used to produce behavior. The driver monitors θ , and when $\theta \in S_b(u)$ no change is necessary. However, when $\theta \notin S_b(u)$, the current behavior is not acceptable and must be switched to a behavior that is appropriate for the circumstances. Given the need to switch, any skill $u' \in S_b(\theta)$ can be employed. An algorithm can be outlined for such task management as follows: If $\theta \in S_b(u)$ then $u' = u$; Else $u' \in S_b(\theta)$. This algorithm can be used to determine when a behavior switch is mandatory; i.e., when θ is such that u is not satisficing then a new skill $u' \neq u$ must be selected. In addition to providing an algorithm for detecting mandatory switches, other useful characteristics of the satisficing approach include the ability to account for hysteresis and other variabilities in observed behavior, and the benefits of a set-based decision mechanism. A complete characterization of satisficing decision theory is beyond the scope of this paper.

3 Implications for ACC Design

It is useful to associate ACC functions with a subset of the human driving skills described in the previous section. In the absence of other traffic, an ACC system regulates speed about a preset value and thereby automates $u = SR$, meaning the enabled skill u is **Speed Regulation**. In the presence of other traffic, an ACC system regulates time headway about a preset value and thereby automates $u = TR$ (**Time headway Regulation**). The transition between these skills, including active braking, is a critical aspect of ACC usability. Two alternative methods for such transitions are of importance: engine braking only and active braking. We argue that active braking is a skill that is distinct from engine braking, where the former is limited to the BA (**Braking Active**) skill and the latter is used in the TR skill.

3.1 Basic Hypothesis

Based on the assumed perspective that ACC systems are designed to safely increase comfort, four factors must be considered to ensure both comfort and safety. First, the dynamic behavior of the ACC should be predictable by drivers (necessary for both comfort and safety). Second, the ACC systems should decrease physical workload without placing unrealistic demands on attention management and human decision-making (even for transient intervals). Third, the transfer of authority between automation and human should be seamless, meaning neither the driver nor the automation should be required to work outside the limits of their operation (i.e., neither should be required to work with a θ for which no $u \in S_b(\theta)$). Fourth, the operational limits of ACC performance should be easily identified.

Regarding these factors, we are primarily interested in when θ is such that the ACC is not satisficing, $u_{ACC} \notin S_b(\theta)$. Such an event can occur if either the ACC system malfunctions or the state of the environment is outside of the scope of the ACC system. Focusing on the second occurrence (we will assume that the first occurrence is negligible — an assumption that must be considered in practice) our task is to determine the perceptual trigger between satisficing ACC behavior and unacceptable ACC behavior. Since the limits of ACC behavior as a function of traffic, weather, time of day, and infrastructure correspond to bounds of $S_b(u_{ACC})$, this task translates into detecting and interpreting the bounds of the satisficing set. Our fundamental hypothesis of effective ACC design is as follows:

Hypothesis 1 *Assuming θ receives attention, switches from ACC to driver control via driver intervention are easiest for the driver if $S_b(u_{ACC}) = S_b(TR) \cup S_b(SR)$.*

In other words, an ACC system that automates both speed and time headway regulation is most likely to facilitate attentionally manageable and seamless transitions from automation to human control. If Hypothesis 1 holds, then the set of existing perceptual cues θ used by the driver to detect mandatory switches to active braking can also be used detect when the ACC system should be disengaged. If Hypothesis 1 is violated, then drivers require either training or a surrogate system for detecting a mandatory switch. Note that this requires knowledge of a driver’s subjective perceptual boundaries beyond which they actively press the brake. As reported in subsequent sections, these boundaries are determined by experiment.

3.2 Alternatives to Hypothesis

In Figures 2(a)-(d) the support for satisficing ACC behavior is shown in relation to the support for TR, SR, and BA behaviors for three idealized cases. Compare each of the cases in Figures 2(a)-(c) to Hypothesis 1 in Figure 2(d) wherein the sets overlap. For Figure 2(a), the ACC system does not accomplish its stated objective of automating car following and speed regulation behaviors and, consequently, is not useful from a designer’s perspective.

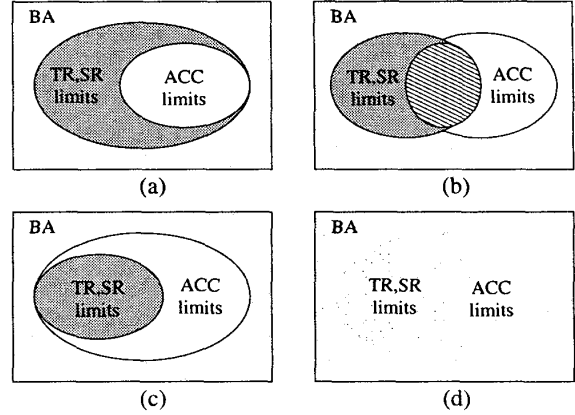


Figure 2: Comparison of TR/SR domains and ACC domain: (a) TR/SR domain broader than ACC domain $S_b(u_{ACC}) \subset S_b(TR) \cup S_b(SR)$, (b) TR/SR domain and ACC domain incompatible $S_b(u_{ACC}) \not\subset S_b(TR) \cup S_b(SR)$, (c) ACC domain broader than TR/SR domain $S_b(u_{ACC}) \supset S_b(TR) \cup S_b(SR)$, and (d) ACC domain approximately equals TR/SR domain $S_b(u_{ACC}) \approx S_b(TR) \cup S_b(SR)$.

For Figure 2(b) the set of states for which ACC and TR/SR are applicable are incompatible. Such a design can either make it difficult for drivers to intervene in ACC control, or require either a period of driver adaptation to learn the limits of the new system or the inclusion of a surrogate system to indicate the limits of the ACC. For a useful system with a wide range of drivers, it is undesirable to design the ACC system that requires the driver to learn and carefully monitor the automation to produce safe automation.

For Figure 2(c), the ACC system exceeds the driver’s capabilities. Such a system appears attractive in that more driver behaviors than just TR and SR are automated. The problem with this approach is that the ACC system does not automate all of the BA skill. By contrast to Figure 2(a) in which a driver knows when to brake because the driver has clearly defined boundaries between TR/SR and BA skills, a driver has no such experience in detecting the limits of ACC behavior. Unless the limits of this behavior are easily perceived by the driver, such a system can result in an unsafe ACC design. For example, consider the results from [5] wherein an ACC system that included a limited amount of braking was studied. Among other things, the study reported that (a) ACC users have “too large [of] expectations” about ACC abilities, (b) most ACC users waited to intervene until a warning was received, and (c) collisions that occurred (in the driving simulator) when the ACC was engaged “could [not] be explained . . . by decreased level of driver alertness” (which may not hold when drivers are not engaged in an experimental study). The first two of these findings suggest that when an ACC system does more than automate TR and SR skills, then drivers have a difficult time safely detecting the limits of the ACC system (there is no natural switch between skills). Instead, drivers sometimes adopt a “wait and see attitude” that allows the

ACC system to reach its limits rather than proactively intervening to avoid an unsafe situation. The last finding suggests that the negative effects of ACC behavior cannot be entirely attributed to not attending to θ .

3.3 Implications from Hypothesis 1

Based on our previous discussion, it is apparent that common sense design supports Hypothesis 1. When hypothesis 1 holds, there are some important implications for ACC design. In the remainder of this section, we discuss a few of these implications.

State Space Selection. Designing a system that satisfies Hypothesis 1 can be considerably easier when both automation and human use the same state space. The reason for this is that the hypothesis depends on the boundary of $S_b(TR) \cup S_b(SR)$ which is most easily determined in the driver's perceptual state space. When the limits of ACC behavior and skilled human behavior match, the human can better interpret system performance and detect the limits of this performance. This helps the driver to naturally switch from ACC behavior to BA behavior via intervention.

Mental Model Switching. In the context provided by SDT, we conclude that a skill switch is mandated when $u \notin S_b(\theta)$. Detecting such a mandatory switch is tantamount to determining if u is in the support for θ . When $S_b(u_{ACC}) = S_b(TR) \cup S_b(SR)$ then it follows that the test for satisficing automation $u_{ACC} \notin S_b(\theta)$ can be performed by the existing tests $TR \notin S_b(\theta)$ and $SR \notin S_b(\theta)$ that are native to the driver. This allows the driver to use existing decision mechanisms to detect the limits of automation without the necessity of excessive training or surrogate assistance.

Surrogates. ACC automates two skills normally performed by automobile drivers: speed regulation and nominal car following. This leaves braking to avoid collisions and other emergency behaviors to the driver. When the behavior of an ACC system exceeds the support for SR/TR skills how do drivers know when to intervene? Provided that drivers can attend to the perceptual cues, two methods for detecting the need to intervene are possible. First, drivers can be trained to learn the perceptual boundary and second, a surrogate can be used to help the drivers detect the boundary. Since the people who drive vehicles come from diverse backgrounds with diverse skills and training, it is unlikely that training will be universally effective. The second option is to use a surrogate to assist drivers in detecting the boundary. This not only helps drivers detect the need to intervene in ACC control, it also acts to train drivers regarding the limits of ACC behavior. However, the design of such a surrogate is a non-trivial task as demonstrated by the difficulty of designing a useful warning system. Such warning systems must be designed with careful attention to driver perceptual and information-processing capabilities. However, a carefully designed surrogate might increase the safety of any ACC system since Hypothesis 1 is an ideal and difficult to reach in practice.

4 Experimental Support for Hypothesis 1

To test Hypothesis 1, we have gathered evidence from the transition between conventional cruise control and active braking. We present this evidence in this section.

4.1 Experiment I

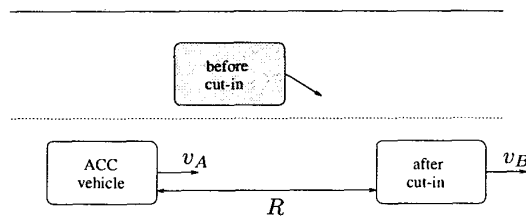


Figure 3: "Cutting in" problem. The lead vehicle prior and subsequent to the cut-in event is represented by a shaded box and an open box, respectively.

To determine SDT-based models of driver behavior, we will focus on the "cutting in" problem wherein vehicle B cuts in front of the driver's vehicle (vehicle A) as diagrammed in Figure 3. Subsequent to a cut-in event, we refer to the vehicle that cuts in as the lead or cut-in vehicle. In the figure, v_A and v_B represent the velocities of the driver's vehicle and the lead vehicle, respectively, $v_R = v_B - v_A$ represents the relative velocity between the vehicles, and R represents the range (relative distance) between the vehicles.

Nissan's SIRCA simulated driving environment includes approximately six miles of highway with three lanes in each direction and ambient traffic. In an experiment using the SIRCA environment, a subject performs lateral control but engages a cruise control (CC) system to perform longitudinal control about a preset condition ($v^* \approx 20\text{m/s} \approx 43\text{mph}$). During the experiment, a cut-in vehicle passes the subject's vehicle while the CC is engaged and cuts into the lane with a specified relative velocity $v_R(0)$ and randomly selected initial time headway $T_h(0)$. Data were partitioned into two classes: *active braking* (brake pedal depressed) and *nominal behavior* (CC engaged, accelerator depressed, or engine braking).



Figure 4: Actual (dashed line) and approximated (solid line) membership functions: (a) accuracy and (b) liability.

Empirical estimates of accuracy and liability can be obtained as described in [3]. Figure 4 presents the resulting empirical estimates and the best fit curve to these estimates.

For the driver to switch from one skill to another, it is necessary to identify when $u \notin S_b(\theta)$. Using $\mu_A(T_c^{-1})$ and

$\mu_L(T_h)$ from Figure 4, we can construct the set of states $S_b = \{\theta : \mu_A(T_c^{-1}) \geq b\mu_L(T_h)\}$ that support nominal behavior, and the set of states $S_b^c = \{\theta : \mu_A(T_c^{-1}) < b\mu_L(T_h)\}$ (superscript c denotes complement) that do not support nominal behavior. If $u \in \{\text{TR}, \text{SR}\}$ and $\theta \in S_b^c$ then $\theta \notin S_b(u)$. Thus, the line $\mu_A(T_c^{-1}) = b\mu_L(T_h)$ determines when behavior must be switched from nominal to braking. In other words, the line is the boundary of $S_b(\text{SR})$.

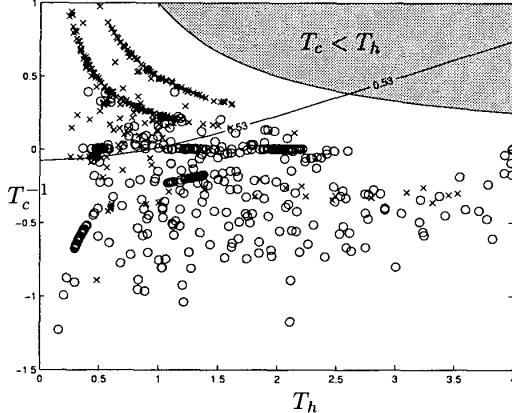


Figure 5: Scatter plot of nominal and braking perceptual states.

Given the empirically derived membership functions, we can determine the boundary between nominal and braking behaviors as a function of b by finding the perceptual states θ for which $\mu_A(T_c^{-1}) = b\mu_L(T_h)$. This is illustrated in Figure 5 for the data gathered in the simulator experiment, where \circ indicates $\theta \in \text{NOM}$ and \times indicates $\theta \in \text{BRK}$. To the northwest of the line, BA is satisficing but TR and SR are not, and to the southeast of the line TR and SR (and, perhaps, BA) are satisficing. Classification can be performed by finding the value of b that optimally separates braking from nominal behavior. Consider the following three performance indices: $J_1(b)$ is the percentage of trials that are incorrectly classified (i.e., the total number of \circ 's above the line plus the total number of \times 's below the line), $J_2(b)$ is the percentage of nominal trials that are incorrectly classified as braking (i.e., number of \circ 's above the line), and $J_3(b)$ is the percentage of braking trials that are incorrectly classified as nominal (i.e., number of \times 's below the line). The value $b = 0.53$ is the minimax value $b = \arg \min_{b \geq 0} \max\{J_1(b), J_2(b), J_3(b)\}$ which attempts to balance the percentage of misclassifications ($J_1(b)$), false alarms ($J_2(b)$), and missed detections ($J_3(b)$). The value $b = 0.20$ minimizes the number of samples misclassified $b = \arg \min_{b \geq 0} J_1(b)$. The classification results for the different values of b are shown in Table 1 and indicate that, on the average, over 85% of samples are correctly classified.

4.2 Experiment II

Because Experiment I relied on a fixed-base driving simulator, there is some question about how these results relate to

b	% misclassified	% false braking	% missed braking
0.20	10.04	1.95	8.09
0.53	13.25	8.37	4.88

Table 1: Classification accuracies for different values of b .

situated driving in real vehicles. To test the transfer of these results to driving, a second experiment was conducted with professional drivers responding to unpredictable cut-in events with real vehicles on a test track.

In the experiment, two vehicles drive in the same lane on a closed test track. The subject drives vehicle A which follows vehicle B. The drivers in vehicles A and B are required to maintain an assigned speed $v_A(0)$ and v_B until a chime rings in vehicle A's car. When the chime rings, the driver of vehicle A is to establish a natural following distance (i.e., drive as if vehicle B had just cut-in to vehicle A's lane) while vehicle B maintains a constant speed. Measurements include R , v_A , brake pressure β , and throttle opening angle α . Time headway and time to collision were computed from these measurements. A complete description of this experiment can be found in [3].

To perform the classification, $\mu_A(T_c^{-1})$ and $\mu_L(T_h)$ were estimated, and the b that minimizes the misclassification error was determined. The results indicate one false alarm (\circ above the line) and no missed detections (\times below the line). The results between the test track experiments and driving simulator experiments are very similar. The test track results produce a slightly smaller value of b ($b = 0.21$ for the test track versus an average value of $b = 0.53$ for the driving simulator) and a slight change in the liability function (achieves its maximum $\mu_L(T_h^{\max})$ at $T_h^{\max} > 0$ for the test track versus $T_h^{\max} < 0$ for the driving simulator). These differences simply indicate that the costs of error are higher when real vehicles are used; in other words, a real collision on the test track is much more costly than a simulated collision in the driving simulator.

5 Conclusions and Future Work

Skilled human driving can be organized into behavioral quanta that correspond to separate skill-based behaviors. For longitudinal control, such skills include following a lead vehicle, braking to avoid a collision, and regulating speed about a desired value. When automation is added to a vehicle, some of these skill-based behaviors are performed automatically by the vehicle itself. We hypothesized that effective human interaction with an ACC system, that is, an interaction resulting in safe and comfortable vehicle dynamics, is achieved when the onset of automated braking matches that of a skilled human operator, and when a human driver can easily detect and interpret the operational limits of ACC-generated braking. By measuring human subjects responses to cut-in events in both driving simulator and test track studies, we have presented experimental support that natural boundaries exist between automated speed regulation (conventional cruise con-

trol) and skill-based braking behaviors. These experimental results are described in the theoretical and computational framework provided by using satisficing decision theory to describe switches between multiple mental models. Extending these results to ACC systems we hypothesize that, assuming an attentive driver, switches from ACC to driver control (via driver intervention) are easiest if the operational limits of ACC behavior correspond to the natural boundaries between that of speed regulation/car-following skill-based behaviors and that of active braking skill-based behavior. We then explored some of the consequences that result when this hypothesis is violated, and observed that appropriately designed surrogates (such as a warning system) can be used to help drivers detect and interpret the operational limits of ACC behaviors. We conclude that advanced vehicle system design can benefit from careful analysis of human interaction with automation by producing systems with an operational domain that is interpretable by human drivers. We further conclude that the analysis of human-vehicle interaction can be systematically performed using a satisficing-based description of multiple mental model dynamics.

References

- [1] J. J. Gibson and L. E. Crooks. A theoretical field-analysis of automobile-driving. *The American Journal of Psychology*, 51(3), 1938.
- [2] M. A. Goodrich and E. R. Boer. Semiotics and mental models: Modeling automobile driver behavior. In *Joint Conference on Science and Technology of Intelligent Systems ISIC/CIRA/ISAS'98 Proceedings*, Gaithersburg, MD, September 1998.
- [3] M. A. Goodrich, E. R. Boer, and H. Inoue. Brake initiation and braking dynamics: A human-centered study of desired ACC characteristics. Technical Report TR-98-5, Cambridge Basic Research, Nissan Research and Development, Inc., Cambridge, MA 02142, USA, 1998.
- [4] M. A. Goodrich, W. C. Stirling, and R. L. Frost. A theory of satisficing decisions and control. *IEEE Transactions on Systems, Man, and Cybernetics — Part A: Systems and Humans*, 28(6):763–779, November 1998.
- [5] L. Nilsson. Safety effects of adaptive cruise controls in critical traffic situations. In *Proceedings of the Steps Forward, Volume III*, pages 1254–1259, Yokohama, Japan, November 9-11, 1995. the Second World Congress on Intelligent Transport Systems.