

Designing Human-Centered Automation: Tradeoffs in Collision Avoidance System Design

Michael A. Goodrich, *Member, IEEE*, and Erwin R. Boer, *Member, IEEE*

Abstract—Technological advances have made plausible the design of automated systems that share responsibility with a human operator. The decision to use automation to assist or replace a human operator in safety-critical tasks must account for not only the technological capabilities of the sensor and control subsystems, but also the autonomy, capabilities, and preferences of the human operator. By their nature, such human-centered automation problems have multiple attributes: an attribute reflecting human goals and capabilities, and an attribute reflecting automation goals and capabilities. Although good theories exist that describe portions of human behavior generation, in the absence of a general theory of human interaction with complex systems, it is difficult to define and find a unique optimal multiattribute resolution to these competing design requirements. We develop a systematic approach to such problems using a multiattribute decomposition of human and automation goals. This paradigm uses both the satisficing decision principle which is unique to two-attribute problems, and the domination principle which is a common manifestation of the optimality principle in multiattribute domains. As applied to human-centered automation in advanced vehicle systems, the decision method identifies performance valuations and compares the safety benefit of a system intervention against the cost to the human operator. By so formulating the problem, the burden of proof is placed on the automation system: *to invoke automation actions, the projected safety-enhancement must be compelling enough to justify the cost to the operator's autonomy.* This effectually integrates human factors considerations into the automation design process from its inception. We illustrate the method by analyzing an automated system to prevent lane departures.

Index Terms—Advanced vehicle systems, decision making, human-centered automation, human factors, satisficing.

I. INTRODUCTION

WITH the rapid progress of technology, there is a concerted effort to use automated systems to augment human abilities in safety critical tasks such as automobile driving, aviation, and process automation. However, a lesson learned from process automation is that, in the absence of human factors considerations, even technologically state-of-the-art systems can be more problematic than beneficial [56], [2], [42]. This lesson indicates the importance of

including human factors in the automation design process so as to prevent “ironies of automation” [2].

Collision and accident avoidance systems (CAASs) include systems that aid in lane keeping, car following, curve negotiation, obstacle avoidance, etc. CAASs are an important component of advanced vehicle control systems and may be realized with minimal or no changes to existing vehicles and highway infrastructure [5], [30]. When coupled with other aspects of intelligent vehicles and highways, CAASs can serve to enhance safety and increase highway throughput. As a result of much current academic and industrial research, the complicated technological and human factors associated with CAAS design are being unraveled thereby enhancing the desire to include CAASs in vehicle design [1], [8], [7], [29]. In essence, designers seek to a) identify situations wherein drivers exhibit unsafe behavior and b) design a CAAS that assumes some responsibility for safety in these situations. In this paper, we formalize human-centered design principles and illustrate their application using an automation system that assists drivers avoid unsafe lane departures.

A. Motivation

CAASs are a type of real-time automated systems that share responsibility with a human operator in safety-critical tasks. Of particular significance is that human automobile drivers have widely varying levels of perceptual abilities, physical skills, and technological understanding. Thus human factors *must* be considered in designing CAASs. To satisfy the demands of a poorly designed and overly sensitive CAAS, a driver may need to increase workload which, in turn, can decrease driver situation awareness, comfort, and even safety [23], [38]. For example, an automated system can be designed to sense possible lane departures using machine vision and warn drivers of impending departures. A poorly designed lane-keeping system might compel drivers to drive with intense concentration on maintaining lane position to prevent warnings from being triggered and, consequently, deflect attention from the task of maintaining a safe speed for traffic and road conditions. Similarly, the driver may rely too much on the system if warnings and interventions are falsely assumed to guarantee safety in all circumstances [38]. Consequently, it is desirable to design for the complete system, which consists of both CAAS technology as well as the human driver. We use the term *human-CAAS* system to emphasize the human driver “in the loop” and design a human-CAAS system which explicitly addresses driver autonomy through human centered design. Such a human-CAAS system should a) reduce mutual adaptation between human and system (driver need not continually relearn the system and readjust driving strategy), b)

Manuscript received March 4, 1999; revised December 20, 1999. The Associate Editor for this paper was Prof. Petros A. Ioannou.

M. A. Goodrich was with Nissan CBR. He is now with the Computer Science Department, Brigham Young University, Provo, UT 84602 USA.

E. R. Boer is with Nissan Research and Development Inc., Nissan Cambridge Basic Research, Cambridge, MA 02142-1406 USA.

Publisher Item Identifier S 1524-9050(00)07346-4.

avoid establishing an unrealistically high level of driver trust in the system (maintain driver “in the loop”) [24]–[27], and c) be applicable to both the attentive and temporarily inattentive¹ driver.

The design solutions to these problems must account for not only the purpose and abilities of automation, but also the purpose and abilities of the driver. Consequently, design of such systems is at the very least a multiattribute problem with (often times) competing goals and (possibly very) different roles. For example, in CAAS operation the human commands the vehicle but the automation may frequently interrupt the driver to suggest or enforce a correction. Although good theories exist that describe portions of human behavior generation [41], in the absence of a general theory of human interaction with complex systems, it is difficult to define and find a unique optimal resolution to such multiattribute human–CAAS problems [61]. This places the design of such systems in the class of ill-formed problems² wherein there is a lack of sufficient information, time, or resources to define or find the optimal solution [37]. Ill-formed problems motivate the search for intelligent solutions, the success of which rests, to some degree, upon the belief that finding the optimal decision is not necessary for making justifiable decisions [9], [20], [21]. We present a formal and systematic design approach which replaces the objective of optimal design with a less ambitious (and perhaps more robust) objective of *avoiding error* [49]. This design paradigm employs the satisficing principle of Simon [45], [47], the domination principle from multiattribute utility theory, and the mathematics of Levi’s epistemic utility theory [31] in a theory called strongly satisficing decision theory (SSDT) [49], [12], [17]. Levi’s theory provides a set-based mathematical formalism necessary to incorporate the notion of satisficing into system design. The resulting design paradigm provides a method for systematically designing human–CAAS systems which justifiably avoid error (e.g., false alarms, discomfort to the driver). This error avoidance perspective and its accompanying decision methodology extend to the design of other real-time systems where humans share responsibility with automation for safety critical tasks.

B. Overview of Related Literature

The concept of satisficing, a decision-making paradigm that differs from the *de facto* paradigm of optimality, was first introduced by Simon [45]–[47]. Many cognitive scientists recognize that insistence on optimality is a misplaced requirement in situations of limited resources and information, and that optimality inadequately describes observed behavior in naturalistic settings [46], [47], [9], [60]. Additionally, the definition of and reliance upon an optimal solution has been questioned by Zadeh [57], [58], and other philosophers, scientists, and researchers concerned with pragmatic decision making [20], [21], [48].³

¹In this paper, we do not address the problem of fully autonomous vehicles wherein drivers can be inattentive. Instead, we restrict attention to problems where other tasks may use or deflect attention.

²Examples of an ill-formed decision problem and an ill-formed control problem are, respectively, the real-time optimal solution to the Traveling Salesman Problem, and the optimal control of a highly nonlinear system with uncertain and time-varying system parameters.

³Also at 1998 NIPS Workshop on Simple Inference Heuristics versus Complex Decision Machines, Breckenridge, CO, Dec. 1998.

Since Simon, many researchers have presented alternative formalisms of the satisficing concept [35], [36], [34], [33], [50], [51] motivated primarily by the desire to make robust decisions in the presence of uncertainty. These developments compare a utility defined over the consequences of a decision to a decision threshold. This decision threshold depends only on observations and not on decision consequences. SSDT is similar to these other developments in that it addresses robustness [19] but, by contrast to comparing a single utility to an action-independent threshold, SSDT compares two utilities defined over the consequences of a decision against each other whence SSDT mathematically generalizes these decision rules (i.e., the decision threshold depends upon both the state of nature and the consequences that result from taking action).

Building on both Simon’s work as well as the notion of bounded rationality, satisficing decision making has seen a recent resurgence of interest in the field of artificial intelligence. This interest is primarily devoted to bounded search methods and has produced some interesting work in anytime algorithms and constrained optimization (see, for example, many papers in [44] or the anytime algorithm tutorial [59]). These and the related economics-based approaches (such as [28]) propose that satisficing emerges from the constraints of bounded rationality and are therefore tantamount to constrained optimization. Satisficing, as we use it, differs from these efforts in that it becomes a formal decision principle in itself instead of either a heuristic determining the structure or termination of a search algorithm or a variant of constrained optimization [32].

Human interaction with automation is an area of current research interest in both academia and industry. Historically, research in human-centered automation has focused on aviation and process automation. Because of the variety of automobile drivers, there are large variances in physical skills, perceptual abilities, technological understanding, etc. These variances make automating these systems somewhat unique. Consequently, it is sometimes difficult to translate research findings from other fields to driving. However, the fields of aviation and process control have some transfer to driving and some of this work has been significant [25], [24], [26], [38], [2], [42], [53].

Two types of automation systems are candidates for inclusion in advanced safety vehicles. We may broadly categorize these as *task automation* and *response automation*. Task automation systems, such as conventional and advanced cruise control, are initiated by the driver with the purpose of relieving some of the physical or mental burden of performing a particular task (such as regulating speed). The difficulty in designing these systems is in helping drivers detect and respond to the behavioral limits of the automation. By contrast, response automation (of which CAASs are a special case) such as lane assistance systems are initiated by the automation to facilitate safer driving. The difficulty in designing these systems is not only to help drivers detect and respond to situations that fall outside their normal range of operation, but also to intervene when necessary to increase safety. The primary purpose of task automation systems is to safely promote comfort, and the primary purpose of response automation systems such as CAASs is to comfortably promote safety. Extensive human factors research has been performed on task automation [55], especially on advanced cruise control [3],

[22], [40]. By contrast, systematic approaches to human-centered response automation have received less attention in the literature [30], [29], [23], [39].

C. Outline of Paper

This paper is divided into three sections. In Section II we identify design principles and the corresponding tradeoffs that arise in the human-centered design of lane assistance systems. In Section III we present a formal theory that allows a designer to systematically address human-centered design for problems where human and automation share responsibility. In Section IV we use observations and measurements to systematically identify design parameters and the resulting decision logic for our example lane assistance system. Finally, we present conclusions from this work.

II. CASE STUDY: CAAS DESIGN PRINCIPLES

It is necessary to design human-CAAS systems which not only make the vehicle more safe, but which also allow the driver to maintain normal attentive vehicle control; such systems bridge the gap between unassisted driving and fully autonomous vehicles. In this section, we design the decision logic for a lane departure CAAS, and highlight important principles in preparation for introducing how SSdT can be used to systematically design CAAS logic.

To systematically design a CAAS, it is useful to consider it as a decision problem. In a decision problem, there are the following elements: the state of the environment Θ , observations of these states X , the decisions⁴ that can be made U , the consequences C that result when a decision is made given the current state of the environment, and values that encode an agent's preference pattern over the set of decision consequences (see Fig. 1). In this context, the CAAS design problem is to: i) identify states of nature (in this case, driving states) that produce unsafe consequences unless rectified, ii) determine what actions can be taken to produce safe consequences, iii) properly assess the actual consequences (including impact on safety and driver autonomy) that result when the CAAS actions are taken, and iv) determine a decision rule that selects an appropriate action for the observed state. Steps two and three must be carefully coordinated so that, in an effort to address step one, the designer does not introduce problems that are worse than they were initially. In the following sections, we discuss states, actions, consequences, values, and action selection for a human-centered lane departure system designed to produce the desirable consequence of increasing system safety without interfering with driver autonomy.

A. Driving State

There are three essential features which affect the consequences of a CAAS action: the state of the driving environment, the state of the vehicle, and the state of the driver. The state of the driving environment includes not only the road profile,

⁴Note that the set of relevant actions should, in practice, be restricted by the driving environment. In Fig. 1, we suppose that the set U contains all possible actions, and then allow the *action selection* module to consider only relevant actions.

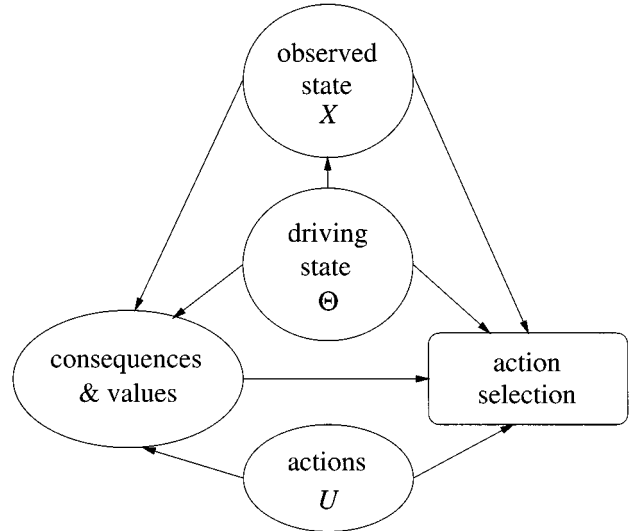


Fig. 1. The design problem.

but also the states (positions, velocities, future positions, future velocities) of other vehicles. The state of the vehicle refers to the vehicle position, velocity, etc., and associated vehicle limitations. The driver's state includes the driver's intentions (turn, stop, change lanes, etc.), conditions (fatigued, intoxicated, alert, irritable, etc.), and limitations (visual acuity, reaction time, etc.). Assuming that the driver will not change current steering angle (methods for inferring driver intention are being developed [43]), these three features are parameterized⁵ by estimates of the Time to Lane Crossing (TLC) which is denoted by τ . Much experimental evidence exists that drivers use TLC to determine when to perform a lane-corrective steering action [10], [11]. We use TLC as the primary decision variable because it can be perceived by drivers and therefore facilitates driver understanding and interaction with the automation [15], [14], [16]. An intuition into the applicability of TLC can be gained from [4] which states, "TLC provides drivers with an estimate of the criticality of the situation and is an upper bound on the time available to turn the vehicle parallel to the road again." Additionally, TLC can be used in computational models that closely emulate skilled driving behavior [4].

Although a number of methods exist for estimating TLC including those reported in [30], [29], for demonstration purposes we use a simple prediction of future lane positions that is discussed in Appendix I-A which is based on an estimate of the current vehicle state. (To reduce false alarm probability in practice, it may be useful to consider more sophisticated estimates of TLC.) This simple TLC estimation scheme consists of three steps: a) estimate the current vehicle state, b) predict the trajectory of the vehicle, and c) estimate the time until lane departure. Formally, when $\hat{y}_r(t + \delta t)$ denotes the estimate of lane position of the vehicle assuming fixed heading (i.e., fixed steering angle) at some future time $t + \delta t$, and when v_w and r_w denote

⁵The TLC estimate used here tells us little about driver state since it is obtained under the assumption that the driver does not change the current steering angle (inattentive driver assumption). If the driver state can be inferred (attentive, preparing to make a lane change, etc.) then this information should be used to calculate a more reasonable TLC.

vehicle and lane width (see Fig. 7(b) in Appendix I-A), respectively, then time to lane crossing is defined as the time before the vehicle's lane position exceeds the lane boundary

$$\hat{\tau} = \min_{\delta t} \left\{ \delta t: |\hat{y}_r(t + \delta t)| > \frac{rw - ww}{2} \right\}. \quad (1)$$

B. CAAS Actions

Given an estimate of TLC, the designer will determine a range of TLC values that indicate an unsafe driving situation; in this case, an impending lane departure. For these unsafe situations it is desirable to have the vehicle act in such a way to increase safety. Based on the work in LeBlanc *et al.* [30], [29], two actions are plausible: appropriately signal the driver of an impending lane departure and thereby help the driver avoid the incident, or intervene in lateral vehicle control in such a way that the departure is avoided and safety preserved. Thus we restrict attention to the set of CAAS actions $U = \{u_W, u_I\}$, where u_W means the vehicle issues a warning to signal the driver, and u_I means the vehicle intervenes and takes (perhaps partial) control of the vehicle. Such a CAAS requires both a system to issue a warning as well as a system to control the vehicle under intervention conditions. The decision problem is to determine which u_i 's to invoke given the estimate $\hat{\tau}$ of the TLC. By their nature, warnings do not directly control driver behavior and, hence, can be classified among the set of hortatory⁶ systems [39] which require explicit consideration of both human and engineering factors. Additionally, it may be desirable to design interventions such that drivers avoid completely surrendering vehicle control to the CAAS, and this desire also requires careful integration of human factors with engineering design.

C. Lane Departure Consequences and Preferences

Identification of a range of TLC values which indicate an unsafe driving situation require the designer to identify the likely consequences of a CAAS action and the preference patterns among these consequences. The set of consequences for a human-CAAS lane departure assistance system can be partitioned into two attributes resulting in two design criteria: a) "taking action in time to prevent road departures for the largest possible set of departure conditions" to increase vehicle safety, while b) "minimizing false warnings and unwanted interventions" to decrease interference with driver autonomy [29, p. 68]. If critical TLC values are misidentified then accidents can occur which could have been prevented or annoying warnings and undesirable interventions can be issued to some drivers. Additionally, if the set of unsafe TLC values is inflexible to individual drivers, the system can be more problematic than beneficial (e.g., some drivers may have to pay excessive attention to the CAAS system to prevent unwanted warnings/interventions, causing an increase in driver workload) [23].

The primary purpose of the lane departure CAAS is to prevent road departures (maximize vehicle safety). However,

⁶Loosely speaking, a hortatory system advises and encourages decision-makers rather than explicitly controls their decisions. Thus in a hortatory system a decision-maker retains autonomy.

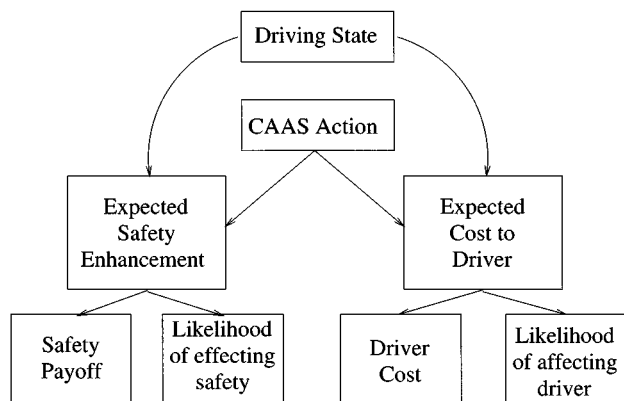


Fig. 2. Decomposition of consequences and values.

since unwanted warnings and interventions incur a cost to the driver (they compromise driver autonomy which can affect driver patience, comfort, attention, and acceptance of the CAAS), counterbalancing considerations should minimize unwarranted warnings and interventions. To account for these two attributes it is useful to create numerical representations of both safety benefit and driver cost. We operationally refer to these numerical representations as *accuracy*, meaning conformity to the fundamental design objective of increasing safety, and *liability*, meaning exposure to the undesirable consequences of impeding driver autonomy, respectively.

We have thus decomposed the consequences of invoking a CAAS action into an accuracy attribute and a liability attribute. The utilities (or, more precisely, utility and inutility) are then represented by numerical accuracy and liability functions. In Appendix I-B, we obtain a two-factor structure of these functions by taking the expected value of a more general function. This structure includes both the *valuation*, which means either safety payoff or driver cost and which is denoted by $J(u; \tau)$, of the CAAS action, as well as the *likelihood*, which means the probability of effecting the corresponding consequence given τ and which is denoted by $\ell(u; \tau)$. The result is the following:

$$\begin{aligned} \text{expected valuation} &= \text{valuation} \times \text{likelihood} \\ \mu(u; \tau) &= J(u; \tau)\ell(u; \tau). \end{aligned} \quad (2)$$

Thus we have not only decomposed the consequences of warning or intervening into accuracy and liability attributes, we have also decomposed these attributes into valuation-likelihood components. This decomposition is beneficial because it facilitates the systematic specification of utility functions. This decomposition is diagrammed in Fig. 2 which corresponds to Fig. 3 from Section IV wherein the functions obtained from measurements are presented.

D. Action Selection

To make a decision, the TLC estimate $\hat{\tau}$, accuracy function, and liability function are determined and passed to the decision logic. We are looking to characterize $\hat{\tau}$ values that justify CAAS actions. The decision logic is based on two principles: a burden of proof concept similar to cost-benefit (liability-accuracy) analysis, and a search through alternative CAAS actions to compare the value of a warning against the value of an

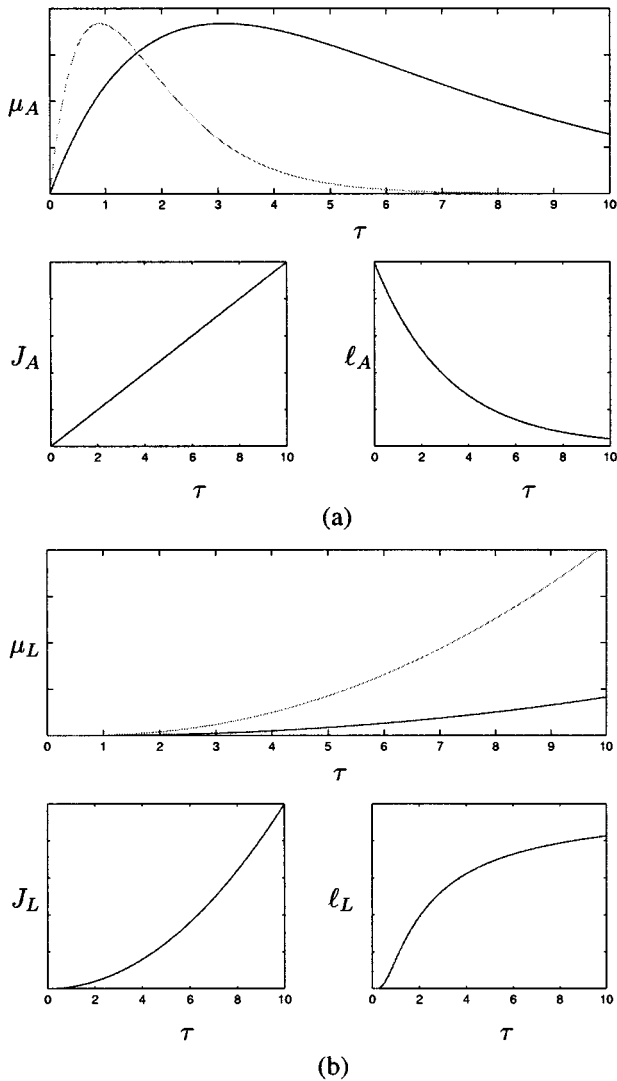


Fig. 3. Expected valuations as a function of decision and TLC—compare Fig. 2. (a) Expected safety enhancement (accuracy) attribute from valuation and likelihood. (b) Expected driver cost (liability) attribute from valuation and likelihood.

intervention. The “burden of proof” concept means that likely safety enhancement must outweigh the likely cost to driver autonomy, and is used to determine which CAAS actions *should* be performed (when proof is sufficient to justify action). In other words, the likely safety enhancement must be compelling enough to justify the likely cost to driver autonomy. The domination concept means that some CAAS actions should not be done because other CAAS actions exist which provide greater safety enhancement with less cost to driver autonomy, and determines which CAAS actions *should not* be performed (because an alternative CAAS action is superior).

III. STRONGLY SATISFICING DECISION THEORY

In the previous section, we introduced principles that influenced CAAS design and operation. From these principles, we distilled the burden of proof and domination concepts as the principles of rational decision making in the tradeoff between

safety and comfort. These principles of rationality can be formally described by satisficing decision theory. In this section, we review a theory of strongly satisficing decisions that places satisficing in a multiattribute, comparison-based mathematical framework that precisely characterizes all decisions that qualify as satisficing. Throughout this section, we use the terms *decision* and *action* interchangeably, though it is more proper to retain the term *decision* to include both *action*, meaning a potential solution to the problem of determining *how to behave*, and *proposition*, meaning a potential solution to the problem of determining *what to believe*.

A. Satisficing Decisions

Simon [45] addressed the issue of limited or bounded rationality by defining an *aspiration level*, such that once this level is met, the corresponding solution is deemed adequate, or *satisficing*. The essence of satisficing is comparison but, by contrast to Simon’s search-based notion, satisficing can be formalized in a multiattribute decision theory wherein attributes are compared to determine justifiable decisions. For the special case of two-attribute problems with one attribute representing the fundamental purpose of seeking a solution and the other attribute representing proximate considerations in reaching this solution, the satisficing decision principle is appropriate.

1) *Practical Decision Making*: Generally, actions cannot be characterized appropriately as being true or false, but may admit functional characterizations instead, such as degree of appropriateness or inappropriateness, expensiveness or inexpensiveness, etc. These characterizations will be determined by the conditions and environment of the decision problem. SSDT was developed as an extension of Levi’s epistemic utility theory [31] to practical problems involving independent benefit and cost attributes [18], [19]. In SSDT, actions are characterized by an *accuracy valuation* μ_A and a *liability valuation* μ_L . To form a systematic design procedure, it is necessary to give operational definitions to characterize these notions.

Accuracy: conformity to a standard. In cognitive contexts the standard is factuality. In practical contexts, the standard corresponds to the fundamental goal or objective relevant to the problem, and accuracy corresponds to the degree of success in achieving that goal.

Liability: susceptibility or exposure to something undesirable. In practical contexts, undesirable consequences may be manifest in the form of costs or other proximate penalties that would accrue simply as a result of taking the action. This cost is quantified in the liability function. Rejecting consequences with high liability improves the set of possible actions.

The fundamental purpose of designing a CAAS is to increase safety; for human–CAAS systems, the safety of the system is the basis for accuracy, and μ_A corresponds to the degree of success the CAAS has in achieving this goal. However, CAASs can interfere with driver autonomy; the cost to driver autonomy is the basis for liability, and μ_L corresponds to the degree to which the CAAS interferes with driver autonomy. By defining these two attributes of the CAAS system, we have encoded both the

TABLE I
SUMMARY OF ACCURACY AND LIABILITY CONCEPTS

Application	Accuracy μ_A	Liability μ_L
Levi's epistemic utility theory	belief	information cost
multi-attribute utility theory (economics)	benefit	cost
CAAS decisions	safety payoff	driver autonomy cost

need to achieve the system's purpose (safety), and the desire to consider the cost of action to the driver (autonomy). In practical terms, the CAAS system should go unnoticed by the attentive driver, but help the inattentive driver. These objectives translate into design criteria of retaining driver autonomy while simultaneously increasing system safety. A summary of concepts related to accuracy and liability, including the human-CAAS system design proposed in this paper, is presented in Table I.

2) *The Satisficing Decision Rule*: Using Levi's error avoidance principle, SSDT provides a method by which the accuracy and liability attributes can be merged: *to avoid error, a decision maker eliminates those decisions which are more liable than accurate*. Although the attributes used in the decision process employ utility-like structures obtained from ordering the consequences of an action, each potential action is evaluated on its own merits without comparing it to other decisions; i.e., actions are not selected as a function of search, but rather as a function of detecting particular environmental conditions. For the lane departure system example, an action is satisficing if it contributes to driver safety more than it interferes with the driver's autonomy. By so formulating the problem, the burden of proof is placed on the automation system: *to invoke a CAAS actions, the predicted safety enhancement must be compelling enough to justify the cost to the driver's autonomy*.

Formally, let U denote the set of possible decisions or actions, and let Θ denote the states of nature. For each decision $u \in U$ and for each state of nature $\theta \in \Theta$, 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).

In SSDT, the set of all decisions which cannot be justifiably eliminated is called the *satisficing set*. The comparative nature of this rule can be identified (using one of the methods of set-valued maximization or fuzzy logic in [18] and [19], respectively) via the following equation:

$$S_b = \{(u; \theta): \mu_A(u; \theta) \geq b\mu_L(u; \theta)\} \quad (3)$$

where b is termed the rejectivity index. Since μ_A and μ_L are constructed independently, the rejectivity serves not only as a relative weighting, but also as a scaling factor which guarantees that the functions representing the two attributes are comparable. (See Section IV-E for a discussion of how b is used to represent driver-dependent preferences for this relative weighting.) Given (3), we can restrict attention to those states of nature which are

⁷For some problems, it is possible, useful, and perhaps essential to describe the consequences of a decision using two attributes, provided a suitable method for merging these attributes can be found. As evidence to this claim, we cite the engineering design examples in [18], [19], the usefulness of cost-benefit analysis, and the importance and success of multi-attribute utility theory in decision analysis [54].

satisficing for a given u , and those controls which are satisficing given the state of nature respectively defined as

$$S_b(u) = \{\theta: \mu_A(u; \theta) \geq b\mu_L(u; \theta)\}$$

$$S_b(\theta) = \{u: \mu_A(u; \theta) \geq b\mu_L(u; \theta)\}.$$

B. Strongly Satisficing Decisions

Although the set S_b contains all possible actions that are legitimate candidates for adoption, they generally will not be equal in overall quality. For example, two satisficing actions may have similar accuracy values (i.e., be equally safe) but have significantly different liability values (i.e., one costs the driver more), and implementing the one with the lower liability will yield essentially the same safety results with lower cost to the driver. Thus we are motivated to further refine the set of satisficing actions. For every $u \in U$ let

$$B_A(u; \theta) = \{v \in U: \mu_L(v; \theta) < \mu_L(u; \theta)$$

$$\text{and } \mu_A(v; \theta) \geq \mu_A(u; \theta)\}$$

$$B_L(u; \theta) = \{v \in U: \mu_L(v; \theta) \leq \mu_L(u; \theta)$$

$$\text{and } \mu_A(v; \theta) > \mu_A(u; \theta)\} \quad (4)$$

and define the set of actions that are *strictly better* than u (i.e., set of actions that dominate u)

$$B(u; \theta) = B_A(u; \theta) \cup B_L(u; \theta) \quad (5)$$

that is, $B(u; \theta)$ consists of all possible actions that are less liable but not less accurate than u , or are more accurate but not more liable than u . If $B(u; \theta) = \emptyset$, then no actions can be preferred to u in both accuracy and liability, and u is a (weakly) *nondominated*⁸ action with respect to θ . The set

$$\mathcal{E}(\theta) = \{u \in U: B(u; \theta) = \emptyset\} \quad (6)$$

contains all nondominated actions. The intersection of this set with the satisficing set yields the *strongly satisficing set*

$$\mathfrak{S}_b(\theta) = \mathcal{E}(\theta) \cap S_b(\theta). \quad (7)$$

From the strongly satisficing set, we can define the support of decision u as those states of nature for which u is strongly satisficing

$$\text{support}_b(u) = \{\theta: u \in \mathfrak{S}_b(\theta)\}. \quad (8)$$

⁸The domination principle and the satisficing principle are independent notions. For some problems, the satisficing principle can be applied without using the domination principle (such as in time-constrained decision making), whereas in multiattribute utility theory only the domination principle is applied. In design problems that employ cost and benefit attributes, we propose incorporating both principles when possible.

This set is the key to systematically identifying those conditions which justify the application of a CAAS action.

C. From Theory to Practice

The theoretical foundation developed in this section allows us to identify the following design steps that can be used in systematically producing human-centered response automation. First, a problem must be identified wherein automation is a candidate for increasing safety. Second, possible solutions to the problem need to be identified. Third, attributes of the consequences of including automation must be identified, including the fundamental attribute of increasing safety and its companion attribute of respecting driver autonomy. Fourth, functions must be identified that have structures which reflect attribute characteristics. Fifth, given these structures parameters must be estimated using careful human factors measurements and engineering analysis of sensor and controller characteristics. Sixth, the resulting accuracy and liability functions should be compared to determine when evidence is sufficient to justify invoking one or more of the CAAS alternatives, and alternatives should be compared to eliminate actions which are dominated by other actions. Given the resulting design, TLC must be estimated and CAAS actions invoked, possibly adapting to driver desires and characteristics.

IV. CASE STUDY CONTINUED: FROM MEASUREMENTS TO VALUES TO ACTIONS

We are now in a position to identify the accuracy and liability functions which, in turn, will be used to determine when warnings should be issued and interventions should be invoked. Identifying these functions is an exercise that includes both human factors and engineering analysis. Selecting appropriate function structures and estimating appropriate function parameters both depend on this analysis. Function selection follows the decomposition process diagrammed in Fig. 2 to produce the functions shown in Fig. 3. In this section, we select function structures that reflect the objective of comfortably promoting safety, and then develop parameter estimates using not only data obtained from the literature, but also a careful analysis and simulation of a particular driver–system interaction. (These measurements are obtained from a model of driver behavior described in [13].) Using the resulting values, we design the decision logic for the lane departure CAAS, and discuss how the safety/comfort tradeoff should reflect individual driver preferences.

A. Determining Values: Choosing the Function Structures

The function structures for the accuracy and liability functions are the same for both warnings and interventions. In this section, we discuss how these structures can be chosen to reflect the objective of comfortably promoting safety. Recall that accuracy and liability are defined completely independently. Consequently, in deriving the accuracy function focus is placed exclusively on safety, and in deriving the liability function focus is placed exclusively on driver autonomy. Also recall the general form (2) where $\mu(u; \tau) = J(u; \tau)\ell(u; \tau)$.

1) *Accuracy*: The accuracy function should embody the design objective of promoting safety. The accuracy valuation is

based on the desire to “prevent road departures for the largest possible set of departure conditions” [29, p. 68]. The function

$$\mu_A(u; \tau) = \alpha(u)\tau e^{-\alpha(u)\tau} \quad (9)$$

which is illustrated in Fig. 3(a), represents the safety enhancement payoff for doing action u when TLC equals τ . The following factors, depicted in their function form in the bottom plots in Fig. 3(a), determine the accuracy.

- **Valuation:** Warnings or interventions are beneficial only if they are issued early enough for the driver/vehicle to respond to them. Therefore, the payoff for a CAAS action increases as TLC increases whence an appropriate monotonically increasing function is required. We use the linear function $J_a(u; \tau) = \tau$ depicted in the bottom left plot of Fig. 3(a) to represent this characteristic because the time available (for the driver in the case of a CAAS warning, and for the controller in the case of a CAAS intervention) to realign the vehicle with the road increases approximately linearly with TLC.
- **Likelihood:** As TLC increases, the relevance of the CAAS action to the immediate vehicle state decreases. This implies that the CAAS action is less likely to effect driver/vehicle corrective behavior and this, in turn, implies that the likely safety enhancement decreases. Therefore, the likelihood of producing the desired safety-enhancing consequence should be represented by monotonically decreasing function. We use the exponentially decaying function $\ell_A(u; \tau) = \alpha(u)e^{-\alpha(u)\tau}$ depicted in the bottom right plot of Fig. 3(a) to represent that the likelihood that a CAAS action will enhance safety decreases exponentially as TLC increases.

2) *Liability*: The liability function should embody the design objective of avoiding interfering with driver autonomy. Clearly, warning a driver or intervening in vehicle control are undesirable if done too early because early CAAS actions imply more “false warnings and unwanted interventions” [29, p. 68] thereby interfering with driver autonomy by demanding attention. The function

$$\mu_L(u; \tau) = \beta(u)\tau^2 F(\tau) \quad (10)$$

which is diagrammed at the top in Fig. 3(b), represents the cost to driver autonomy for taking action u when TLC equals τ . The following factors, depicted at the bottom of Fig. 3(b), determine the liability.

- **Valuation:** A standard quadratic cost function is chosen to represent the cost of early warning and intervention independent of lane departure concerns [23]. The quadratic cost indicates that the cost of interfering with driver behavior rapidly increases with TLC due to the fact that actions taken long before a lane departure will occur much more often and be less helpful to the driver. This indicates that early CAAS actions (those issued at large TLC values) interfere with driver performance more than later actions (those issued at small TLC values). The term $J_L(u; \tau) = \beta(u)\tau^2$ depicted in the bottom left plot of Fig. 3(b) represents this factor.

TABLE II
ACCURACY AND LIABILITY VALUES.

expected valuation	valuation	likelihood	warning	intervention
$\mu_A(u; \tau) = \tau \alpha(u) e^{-\alpha(u)\tau}$	τ	$\alpha(u)^{-\alpha(u)\tau}$	$\alpha(u_W) = \frac{1}{3.4}$	$\alpha(u_I) = \frac{1}{1.2}$
$\mu_L(u; \tau) = \beta(u)\tau^2 \bar{F}(\tau)$	$\beta(u)\tau^2$	$\bar{F}(\tau)$	$\beta(u_W) = 1$	$\beta(u_I) = 5$

- **Likelihood:** The likelihood of interfering with driver autonomy requires information about false alarms. The likelihood depends on the CAAS system false-alarm rate for the attentive driver whence $\ell_L(u; \tau) = F(\tau)$ where $F(\tau)$ is the likelihood of false alarm. Intuitively, because predictions of future driving state and driver behavior become more uncertain as the prediction horizon increases, the probability of falsely predicting a lane departure and thereby interfering with driver autonomy approaches certainty as TLC increases whence $\bar{F}(\tau)$ approaches unity. This is illustrated in the bottom right plot in Fig. 3(b).

B. Determining Values Continued: Estimating Function Parameters

In this section, we discuss parameter selection for accuracy and liability functions for both warnings and interventions. We identify parameters using not only observations reported in the literature, but also careful analysis and simulation of a model of driver behavior and a description of a CAAS intervention controller. Because we wish to allow for a decision logic which can account for various drivers, the parameters should be chosen as general as possible. The accuracy and liability parameters obtained from careful simulations and analysis are summarized in Table II.

An important issue that deserves attention is the variation of the parameters α and β because values and, consequently, decision thresholds are determined by these parameters. The principal functions of the α parameter are to determine not only the timing when a warning/intervention is maximally accurate but also the relative accuracy of such behaviors. The principal function of the β parameter is to determine the relative liability of warnings versus interventions. Sensitivity to these parameter values can only be determined by designing a CAAS and testing the system with human subjects (we illustrate how such tests can be used to select α and β in our simulation study). We thus leave this area for future research, but note that some information about sensitivity can be obtained by comparing the parameter values distilled from the human factors literature to those values obtained from the simulation study.

1) *Accuracy of Warnings:* The accuracy of a warning should be maximum at that value of τ which gives the driver time to estimate the current state and control the vehicle to the lane center. The key to determining $\alpha(u)$ is to note that the maximum value of μ_A occurs when $\tau = (1/\alpha(u))$. From the literature, Evans indicates that most drivers respond to an unexpected driving situation within an interval of 1.5 and 4.0 s, with an average value of 2.5 s [6, p. 121]. The value of $\alpha(u_W) = 1/4$ (implying $\tau_W^{\max} = 4.0$) is chosen as a rough estimate because the warning will best accomplish its purpose (for most drivers) if signaled at 4 s before lane crossing. Thus independent of liability, the ac-

curacy $\mu_A(u_W; \tau)$ (represented by the solid line in Fig. 3(a)) achieves its maximum at $\tau_W^{\max} = 4.0$.

Using careful simulation and analysis, the likely response characteristics of a driver can be determined by the time required to estimate the current lane position, denoted t_O (O for observation), and the time to steer the vehicle to the lane center, denoted t_C (C for control). These characteristics determine the decay rate $\alpha(u)$ and, consequently, the maximum of $\mu_A(u; \tau)$. The control time t_C is defined⁹ as the time required to steer the vehicle from an initial state at the lane boundary $|y_r(0)| = 1.1$ (according to the road and car dimensions in Appendix I-A) to the center of the lane $|y_r(t_C)| = 0$ in the absence of process or measurement noise. Using simulations presented in [13], a first-order estimate of $t_C = 1.2$ s is obtained. The observer time t_O is defined as the time required for the observer to converge from an initial estimation error equal to the lane crossing distance $|\hat{y}_r(0) - y_r(0)| = 1.1$ to zero estimation error $|\hat{y}_r(t_O) - y_r(t_O)| = 0$ in the absence of control (i.e., $y_r(t) = y_r(0)$ for all time). Using simulations presented in [13], a first-order estimate of $t_O = 2.2$ s is obtained. Example time histories of these tasks, obtained from a model of driver behavior [13], are shown in Fig. 4. Accuracy is maximized when the decay parameter is set to $\alpha(u_W) = 1/(t_C + t_O)$. In words, a warning is most accurate if it gives the driver t_O seconds to estimate the state of the vehicle and t_C seconds to control the vehicle.

2) *Accuracy of Interventions:* The intervention accuracy should be maximum at that value of τ which gives the CAAS controller time to control the vehicle to the lane center whence, for intervention, the same factors determine the shape of $\mu_A(u_I; \tau)$ [represented by the dashed line in Fig. 3(a)] but τ_I^{\max} and, hence $\alpha(u_I)$, must be determined. Two constraints dictated by human factors influence the value of τ_I^{\max} . First, an intervention should not occur until some time after the warning occurs, giving a driver a chance to react to the warning. Second, an intervention should occur early enough to allow the intervention controller to smoothly and safely intervene (actuator time constant). From the literature, a rough estimate of $\alpha(u_I) = 1/2$ (implying $\tau_I^{\max} = 2$ s) indicates that the intervention will best accomplish its purpose if applied at 2 s before lane crossing and 2 s after warning τ_W^{\max} .

Using careful simulation and analysis, the intervention accuracy should be maximum at that value of τ which gives the CAAS controller time to control the vehicle. Since the CAAS controller is fully attentive, only controller time is relevant ($t_O = 0$ since the sensors continually monitor the road). This yields a decay parameter of $\alpha(u_I) = 1/t_C$, which gives maximum intervention accuracy at $t_C = 1.2$, where we have bounded the performance of the intervention controller by the

⁹Alternatively, t_C can be defined as the time it takes for the driver to cause TLC to reach infinity (no movement toward lane boundary). This measurement indicates that the vehicle is again under attentive driver control and assumes that the driver is not engaged in unstable steering oscillations.

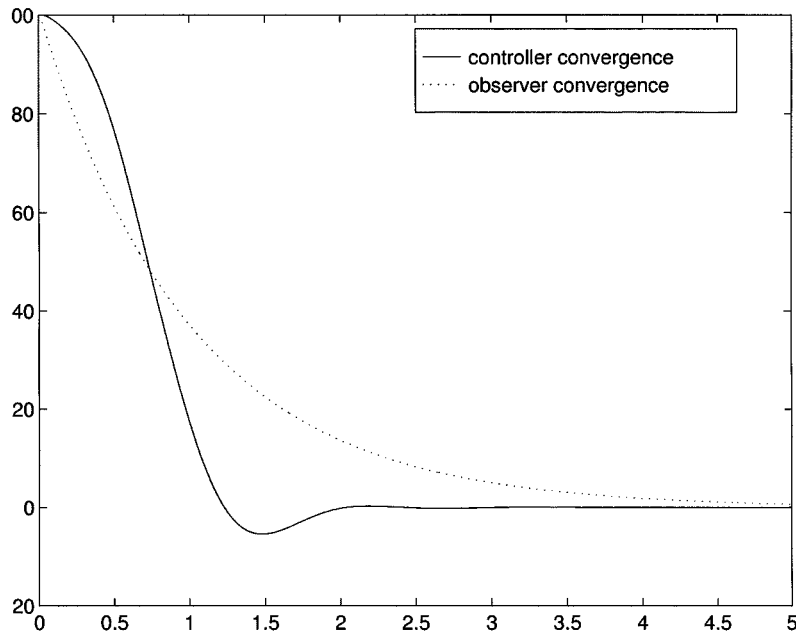


Fig. 4. Convergence times: time to control vehicle from lane edge (100% error) to lane center, and time to observe true vehicle state from lane width error (100% error) to zero. Data were obtained from a model of skilled driving behavior.

estimates of driver behaviors (the CAAS controller should be at least as responsive as a driver).

3) *Liability of Warnings*: From the literature, the nominal value $\beta(u_W) = 0.2/e$ yields a warning threshold consistent (for $b = 1$, see Section IV-E) with those derived from the subjective driver preferences reported in [30], [29]. Note that this parameter is approximately driver-independent since it is determined for the “average” driver.

Using careful simulation and analysis, we can set $\beta(u_W)$ equal to unity since only the ratio $(\beta(u_I)/\beta(u_W))$ needs to be specified.¹⁰ The false alarm probability is defined as the likelihood of predicting $|\hat{y}_r(t+\tau)| > 1.1$ (and thus invoking a CAAS action) when in fact $|y_r(t+\tau)| \leq 1.1$ (and thus no CAAS action is warranted). This yields the cumulative distribution function

$$F(\tau) = \int_0^\tau P(|\hat{y}_r(t+\rho)| > 1.1 \text{ and } |y_r(t+\rho)| \leq 1.1) d\rho.$$

In the bottom right plot of Fig. 3(b), the line represents the empirical cumulative distribution function for false TLC prediction under the conditions described in [13].

4) *Liability of Interventions*: Clearly, $\beta(u_I) = N \times \beta(u_W)$ sets the cost of an intervention N times higher than the cost of a warning. From the literature, $N = 4$ produces an intervention threshold that agrees with those derived from the subjective driver preferences reported in [30], [29]. Careful measurement of N must be determined by experiment with human subjects. Such an experiment was beyond the capabilities of the driver model, and performing an appropriate experiment with human subjects is beyond the scope of this paper. For the results reported using careful simulation and analysis, we subjectively set $N = 5$.

¹⁰To see this, observe that for $b' = b\beta(u_W)$ then $S_{b'} = S_b$. Thus we set $\beta(u_W)$ equal to unity, and determine $\beta(u_I)$ in relation to unity.

C. Action Selection

To make a decision, accuracy, liability, and the TLC estimate $\hat{\tau}$ are determined and passed to the decision logic. The decision logic is based on two principles: the burden of proof concept (i.e., the satisficing principle), and the domination principle. In words, we use the “burden of proof” principle to determine when CAAS actions *should* be performed (when proof is sufficient to justify action), and the domination principle to determine when CAAS actions *should not* be performed (because an alternative CAAS action is superior). The burden of proof can depend on the individual driver preferences whereas superiority depends on system design. As will be shown in the next section, this is a desirable property for systems which should be adapted to individual driver preferences (or, in the more general automation case, specific operator preferences).

1) *Burden of Proof*: To determine the thresholds τ'_W and τ'_I below which warnings and interventions are respectively issued, we must compare the accuracy against the liability. When accuracy exceeds liability then safety interests outweigh costs to driver autonomy and a CAAS action should occur. To perform such a comparison, we must ensure that the accuracy and liability valuations are comparable. We do this via the parameter b that allows us to adjust the relative value of safety to autonomy and thereby adapt to individual differences. The critical TLC values (defined as those values that delineate the region between satisficing and nonsatisficing CAAS actions) occur when accuracy and liability are equal, which happens when the accuracy and liability functions (see, for example, the diagrams in Fig. 5 which represent the accuracy and liability functions with parameters selected using observations reported in the literature) intersect

$$\tau'_W = \arg_{\tau} \{ \mu_A(u_W; \tau) = b\mu_L(u_W; \tau) \} \quad (11)$$

$$\tau'_I = \arg_{\tau} \{ \mu_A(u_I; \tau) = b\mu_L(u_I; \tau) \}. \quad (12)$$

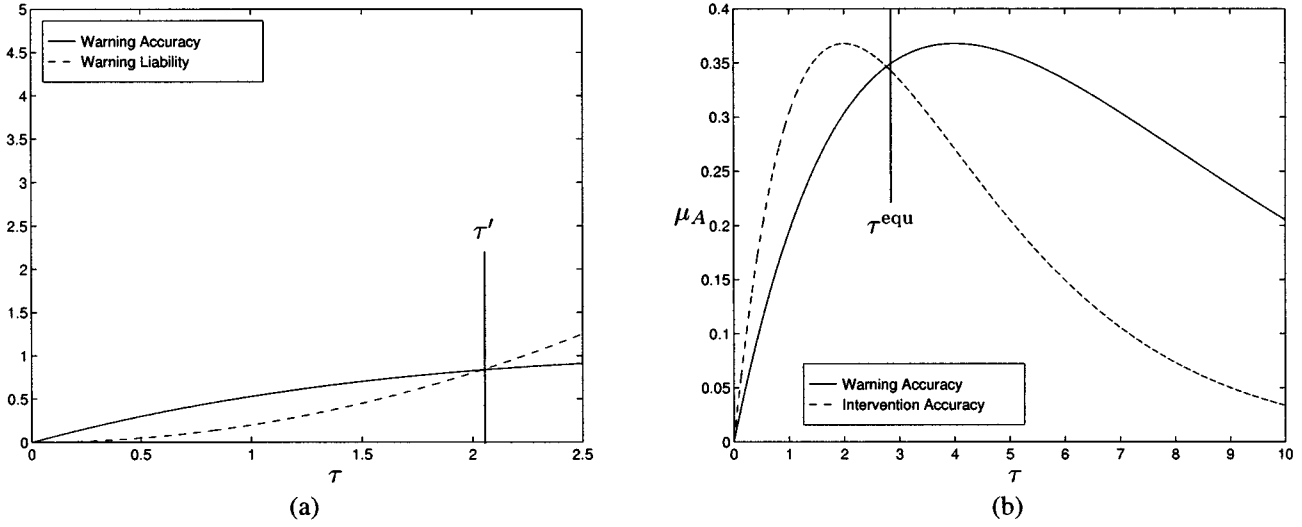


Fig. 5. Application of decision principles. (a) Satisficing, or burden of proof, implies that a warning u_W is justified only if $\mu_A \geq b\mu_L$ which occurs when $\tau \leq \tau'$, and (b) domination implies that u_I is justified only if $\mu_A(u_I; \tau) \geq \mu_A(u_W; \tau)$ (in this case, $\mu_R(u_I; \tau) > \mu_R(u_W; \tau)$ for all τ whence u_I is not dominated by u_W only when $\mu_A(u_I; \tau) \geq \mu_A(u_W; \tau)$) which occurs when $\tau < \tau^{\text{equ}}$. The function parameters in these figures are obtained from observations reported in the literature.

For $\tau \leq \tau'_W$ (to the left of the intersection value), accuracy exceeds liability and, hence, the warning action should be invoked. For $\tau > \tau'_W$ (to the right of the intersection value), liability exceeds accuracy and, hence, the warning action should not be invoked. A similar argument holds for interventions in relation to τ'_I . In Fig. 5(a), the accuracy of u_W exceeds the liability for $\tau \leq \tau'_W = 2.03$ s. Thus the region of support for warning and intervention is delineated by

$$S_b(u_W) = \{\tau: \tau \leq \tau'_W\}$$

$$S_b(u_I) = \{\tau: \tau \leq \tau'_I\}.$$

The accuracy of u_I exceeds the liability for all $\tau \leq \tau'_I = 1.01$ s. Note that for many drivers, $\tau'_W = 2.03$ s is within the range of normal operation.

2) *Eliminating Inferior Alternatives*: Since, in general, warnings will always be less invasive than interventions, attention should be restricted to liability functions defined such that $\mu_L(u_I; \tau) > \mu_L(u_W; \tau)$ for all $\tau > 0$. We ensured this by setting $\beta(u_I) = N\beta(u_W)$ for $N > 1$. Additionally, warnings more effectively promote safety for large TLC whereas interventions more effectively promote safety for small TLC. Clearly, if there are TLC measurements for which warnings not only promote safety more than interventions, but are also less invasive than interventions, then interventions are not useful. The critical value (defined as the value that delineates the range of nondominated CAAS actions) τ^{equ} can be identified as that value for which $\mu_A(u_W; \tau)$ intersects $\mu_A(u_I; \tau)$. For the accuracy and liability valuations shown in Fig. 3, this critical value occurs at the intersection of the two accuracy functions

$$\tau^{\text{equ}} = \arg_{\tau} \{\mu_A(u_I; \tau) = \mu_A(u_W; \tau)\}. \quad (13)$$

For $\tau < \tau^{\text{equ}}$ both warning and intervention actions are permissible since neither CAAS action is obviously superior, but for $\tau \geq \tau^{\text{equ}}$ only warning is permissible since it has both lower

liability and higher accuracy than does intervention (i.e., an intervention is dominated by a warning) whence

$$\mathcal{E}(u_W) = \{\tau \geq 0\}$$

$$\mathcal{E}(u_I) = \{\tau \leq \tau^{\text{equ}}\}.$$

D. Critical Thresholds

Critical thresholds are defined via the strongly satisfying set $S_b \cap \mathcal{E}$ using (11)–(13) as

$$\tau_W = \tau'_W \quad (14)$$

$$\tau_I = \min(\tau^{\text{equ}}, \tau'_I) \quad (15)$$

and the regions of support are defined as those TLC measurements which invoke a CAAS response, and are given by

$$\text{support}(u_W) = \{\tau: \tau \leq \tau_W\}$$

$$\text{support}(u_I) = \{\tau: \tau \leq \tau_I\}.$$

Results for parameters obtained from the literature are shown in Fig. 6(a), where τ_W and τ_I , represented by the solid and dashed lines, respectively, are plotted as functions of b . When $b = 1$, the resulting decision thresholds agree with those reported in [30], [29]. The two horizontal dotted lines indicate the two critical values of τ for a nominal value of $b = 1$, which is represented by the vertical dotted line. For this value of b , the resulting values for τ_W and τ_I coincide with those values proposed in [30], [29]. Observe that the τ_I curve becomes constant for values less than about $b = 0.15$. This occurs because warnings dominate interventions for $\tau \geq \tau^{\text{equ}}$.

Given the accuracy and liability functions obtained from careful analysis and simulation of a driver model, the critical values of τ are identified and plotted as a function of b in Fig. 6(b). Note that intervention occurs only when the time to lane crossing is less than 1.57 s. Thus no interventions can occur unless the system predicts, via (16) and (17) in Appendix I-A, that the driver will leave the lane within about one-and-a-half seconds.

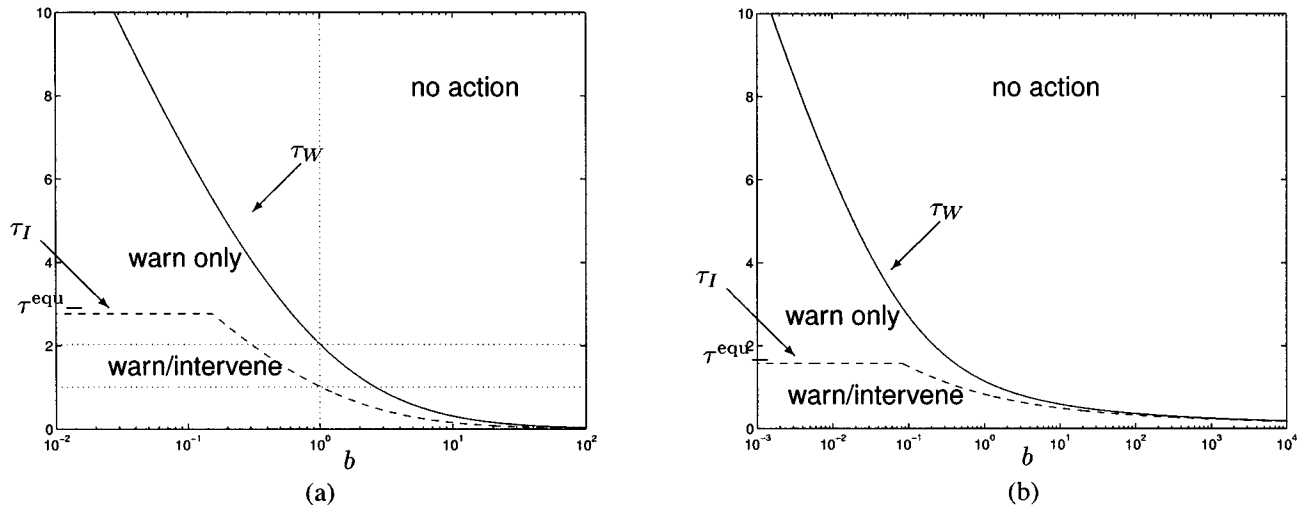


Fig. 6. Critical times for warning and intervention as a function of b for (a) rough estimates and (b) driver model data. The b parameter is a driver-dependent measure of the subjective disposition toward accepting assistance from the warning/intervention system.

E. Driver Dependence

The accuracy and liability functions have been identified in a careful analysis presented in the previous sections. This analysis has assumed that all drivers exhibit a similar value structure. However, it is desirable to allow the resulting decision thresholds to depend upon individual driver preferences. This can be done by adjusting the relative importance of comfort and safety by adapting b and thereby allow for individual differences in the way drivers interact with a system such as a CAAS.

A nominal value for b can be experimentally determined as follows. For any value of b there exist critical thresholds τ_I and τ_W , and for each τ_I and τ_W there is a probability of falsely predicting a lane departure as specified by $F(\tau_I)$ and $F(\tau_W)$. By specifying an acceptable false-alarm rate, a corresponding rejectivity can be determined. For example, based on the driver model, a rejectivity of $b = 10/e$ yields $\tau_W = 0.76$ s and $\tau_I = 0.61$ s, corresponding a 6.9% chance of falsely warning an attentive driver and a 3.7% chance of falsely intervening, respectively.¹¹ Thus the likelihood that an attentive driver will receive a false CAAS action is small, but not negligible. The likelihood of false alarm can be reduced, but at the expense of decreasingly smaller warning and intervention times.

Ideally, the decision logic should not only allow for variations in (and permit adaptation to) individual driver preferences, but should also adapt to changing situations (when, for example, drivers are navigating a sequence of difficult curves wherein they accept TLC values lower than those acceptable in nominal straight-lane driving) and allow for situation-adaptive autonomy [25], [24], [26], [27]. This can be done by adjusting the rejectivity b as a function of the driver's disposition or judgment. If there are drivers for whom a nominal b value is inappropriate (which may be the case when, for example, typical environmental conditions are substantially different from those which determined the false-alarm rate and, therefore, the nom-

inal value of b), then the experimentally determined nominal b value can be (slightly) adjusted to conform to their individual preferences. At one extreme, drivers who do not want the CAAS system to be a factor can be accommodated by letting $b \rightarrow \infty$. For these drivers, the strongly satisficing set is empty unless $\tau = 0.0$. At the other extreme, drivers who, for the sake of safety, do not mind frequent interventions or continuous warnings can be accommodated by letting $b \rightarrow 0.0$. These drivers receive a constant warning (since $\tau_W = \infty$), but no intervention unless $\tau \leq \tau^{\text{equi}}$ (since $\tau_I = \tau^{\text{equi}}$). A similar reasoning can be employed to adjust the system to changing driving conditions (such as sensor noise, weather, traffic density, narrow road shoulder width, etc.), or to adjust the system to important operator-specific tolerances.

Two observations are worth noting: a) because the decision rule is set-valued, it is possible that both warning and intervention can occur simultaneously and b) there is a graded sequence of CAAS actions which increase in severity as the observed TLC decreases (a warning is less severe than an intervention). These observations are consistent with the design principles advocated in [23] which call for a gradual increase in the severity of warnings (which is a form of the coordination of multiple warnings). Potentially, these observations are compatible with the desire to develop a system which adapts to changing environments, including situation-adaptive autonomy. The problems of how and when to adapt the CAAS system and which measurements are most effective in guiding adaptation are important areas of future research.

V. CONCLUSIONS

In this paper, we described a plausible design process to produce human-centered collision and accident avoidance systems (CAASs). Within this design process, both safety-enhancing attributes as well as autonomy-inhibiting attributes must be identified (either explicitly, as we have done, or implicitly using sound "engineering judgment"). These attributes should be compared

¹¹These percentages appear high, but should be interpreted against the base rate of times when TLC drops below critical thresholds for a real driver rather than for the driver model.

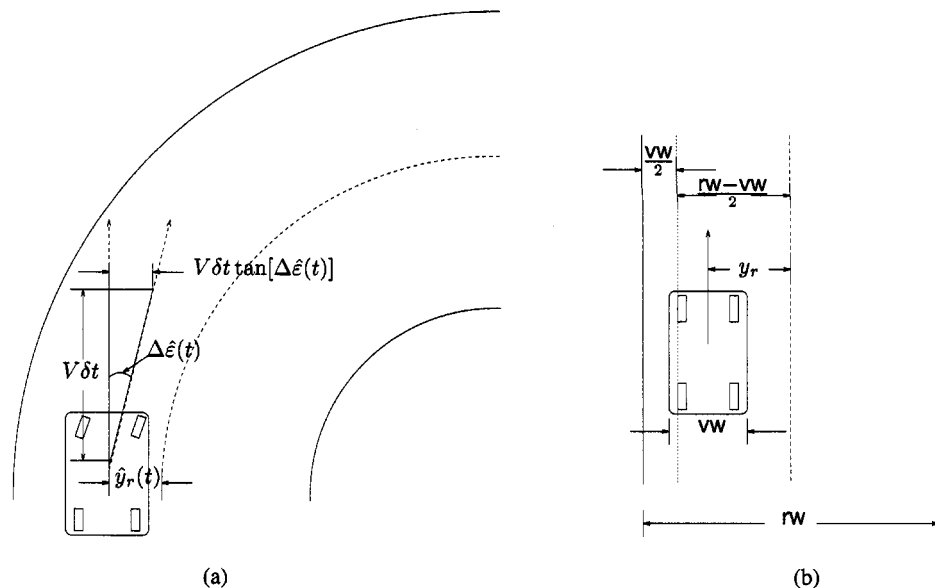


Fig. 7. Predicting TLC. (a) Predicted position ($\Delta\hat{\varepsilon}(t)$) is measured counterclockwise from vertical, and (b) lane departure position. In the figure, the relative widths of the vehicle and the road have been distorted to facilitate the definition of parameters; the dashed line denotes the center of the lane and should not be confused with the dashed line which separates lanes on real two-lane roads.

to tradeoff between these two competing design requirements and thereby produce a CAAS that comfortably promotes safety.

It appears that in systems designed to comfortably increase operator safety, tradeoffs are frequently handled by setting intuitively obtained thresholds. Fortunately, this intuitive design approach can be formalized and, consequently, systematized using strongly satisficing decision theory. Rather than arbitrarily aggregating safety-enhancing and autonomy-inhibiting design criteria into a single performance valuation and then extremizing this valuation, these two criteria should instead be translated into two independent numerical attributes which are compared to determine when CAAS actions are appropriate. This comparison places the “burden of proof” on the automation system, and requires that evidence be compelling enough to justify the cost to operator autonomy. This comparison and the selection of dominating actions determine which set of CAAS actions are appropriate for the observed environment. The flexibility of the resulting design methodology motivates future research efforts to a) apply the principles of situation-adaptive autonomy to such systems, b) coordinate multiple automated actions; i.e., determine conditions which describe when an action can be safely combined with other actions, and c) continue exploring theory-based models of human driving behavior that include how humans react to the introduction of automated systems [15].

APPENDIX I

A. Model of Vehicle Dynamics

Assuming that the driver does not change steering angle, a simple TLC estimation scheme can be performed in the three following steps.

- The current state of the vehicle can be estimated using an observer operating on the vehicle system described by a dynamics equation with state (see Fig. 7)

$$x = [y_r, \dot{y}_r, \Delta\varepsilon, \Delta\dot{\varepsilon}, \delta, \kappa]^T$$

and the elements of this vector are defined as follows: y_r is the lateral distance between the vehicle center of gravity and the center of the lane, $\Delta\varepsilon = \varepsilon - \varepsilon_d$, ε is the yaw angle of the vehicle body, ε_d is the desired yaw angle set by the road, $u = \dot{\delta}$, δ is the front wheel steering angle, κ is the road curvature, and \dot{x} denotes time differentiation. A description of a useful dynamics model and corresponding optimal observer is subsequently provided.

- The predicted lateral position of the vehicle at time $t + \delta t$ can be obtained by assuming that the driver maintains current steering angle and by using the current estimates $\hat{y}_r(t)$ and $\Delta\hat{\varepsilon}(t)$ as follows¹² [Fig. 7(a)]:

$$\hat{y}_r(t + \delta t) = \hat{y}_r(t) + V\delta t \tan[\Delta\hat{\varepsilon}(t)]. \quad (16)$$

- The smallest δt such that $|y_r(t + \delta t)|$ exceeds a fixed value (the lane boundary) is the TLC estimate $\hat{\tau}$. In Fig. 7(b) the vehicle and road are diagrammed. The dashed line represents the lane center, and the distance from this lane center to the center of the car is the lateral lane position y_r . The dotted line is plotted at a distance of one-half of a vehicle width from the lane boundary. If the center of the car crosses the dotted line, then the edge of the car leaves the lane. Thus for a road width of $rw = 3.65$ m, a portion of a $vw = 1.45$ m wide vehicle has departed the lane if its

¹²This estimate holds for sufficiently small δt . For the simulations presented in [13], $V\delta t = 0.25$ m which is much smaller than the average curve length of 25 m.

lateral deviation exceeds $(rw - vw/2) = 1.1$ m. Formally, the time to lane crossing $\hat{\tau}$ is defined as

$$\hat{\tau} = \min_{\delta\tau} \{ \delta t: |\hat{y}_r(t + \delta t)| > 1.1 \}. \quad (17)$$

The dynamics of steering can be estimated by a linear system given by

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx \end{aligned} \quad (18)$$

where $x = [y_r, \dot{y}_r, \Delta\varepsilon, \Delta\dot{\varepsilon}, \delta, \kappa]^T$, and the elements of this vector are defined as follows: y_r is the lateral distance between the vehicle center of gravity and the center of the lane, $\Delta\varepsilon = \varepsilon - \varepsilon_d$, ε is the yaw angle of the vehicle body, ε_d is the desired yaw angle set by the road, $u = \dot{\delta}$, δ is the front wheel steering angle, κ is the road curvature, and \dot{x} denotes time differentiation. The vehicle parameters are given by

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{A_1}{V} & -A_1 & \frac{A_2}{V} & B_1 & A_1 - V^2 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & \frac{A_3}{V} & -A_3 & \frac{A_4}{V} & B_2 & A_4 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

where V is the vehicle speed and where the A_i 's and B_i 's are defined as follows:

$$\begin{aligned} A_1 &= \frac{-2(C_{sf} + C_{sr})}{m} & A_2 &= \frac{2(C_{sr}l_r + C_{sf}l_f)}{m} \\ A_3 &= \frac{2(C_{sr}l_r + C_{sf}l_f)}{I_z} & A_4 &= \frac{-2(C_{sf}l_f^2 + C_{sr}l_r^2)}{I_z} \\ B_1 &= \frac{2C_{sf}}{m} & B_2 &= \frac{2l_f C_{sf}}{I} z \end{aligned}$$

l_f and l_r are the distance from center of gravity to the front and rear axles, respectively, C_{sf} and C_{sr} are the cornering stiffness of the front and rear tires, respectively, and where the output y is observed from the state vector x via the mapping

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

The numerical values used in the simulation were obtained from [52], and velocity was set to $V = 25$ m/s.

B. Mathematical Structure of the Decision Problem

In this appendix, we develop the necessary decision theoretic foundation to justify our derivation of accuracy and liability evaluations. In designing CAAS systems, we are concerned with how actions will affect safety and driver autonomy. Thus we restrict our attention to a two-attribute consequence set $C = C^d \times C^s$ where C^d represents the consequence to driver autonomy and C^s represents the consequence to safety. The automated safety-enhancing system attempts to minimize the cost to operator autonomy while simultaneously maximizing safety. We associated $\mu_A(u, \theta)$ with the consequences C^s wherein we

value a CAAS action because it promotes safety. However, considerations of driver autonomy induce us to consider the cost of invoking a CAAS action (or, equivalently, the benefit of not invoking a CAAS action). Thus we associate $\mu_L(u, \theta)$ with the consequences C^d wherein we value CAAS inaction because it interferes with driver autonomy (the payoff for correct inaction).

1) *Accuracy*: From a strictly safety-enhancement perspective, whenever inaction leads to a lane departure then invoking a CAAS action is correct, and whenever action causes a lane departure then the action is incorrect. The designer must determine the consequences and probability of incorrect actions. If the impact and the probability is high then no CAAS action will ever be justified. Clearly, the impact of invoking a CAAS action that will cause a lane departure (taking incorrect action where inaction would have been safe) is significant. However, a reliable CAAS action will never cause a lane departure (e.g., a warning will never cause the vehicle to swerve off the road, and an intervention will never cause a lane departure) except through the mitigating consequence of interfering with driver autonomy. Thus any cost for falsely invoking a CAAS action is appropriately assigned to the cost for interfering with driver autonomy. Thus the payoff for taking action can be entirely determined by the safety benefit that will result; i.e., any CAAS action has nonnegative safety payoff.

The precise utility of the action depends on a) the probability that the action will produce nominal consequences (no lane departure) given the current driving state and b) the safety payoff that accrues assuming that the action produces nominal consequences. Recall that a consequence is a function of action and driving state $c(u, \theta)$. The probability that action produces safe (nominal) consequences $c(u, \theta) \in N$ given the current state is given by $p(c(u, \theta) \in N | \theta) = p_N(u | \theta)$ where N is the random variable

$$N = \begin{cases} u, & \text{if } c(u, \theta) \in N \\ 0, & \text{otherwise.} \end{cases}$$

The payoff of an action is the increase in safety that results from taking action, and this is denoted by $J_A(u; \theta)$. Combining these factors gives

$$\mu_A(u, \theta) = J_A(u; \theta) p_N(u | \theta). \quad (19)$$

Taking the expected value of (19) with respect to the probability of state θ given observation x and assuming, for simplicity, discrete states, yields

$$\mu_A(u; x) = \sum_{\theta \in \Theta} J_A(u; \theta) p_N(u | \theta) p(\theta | x). \quad (20)$$

Within the scope of this paper, we assume that $p(\theta | x) = \delta(\theta - x)$ where δ is a delta function; this is equivalent to assuming that our observation x precisely determines θ . Given this assumption, (20) becomes

$$\begin{aligned} \mu_A(u; x) &= J_A(u; x) p_N(u | x) \\ &= J_A(u; \tau) \ell_A(u; \tau) \end{aligned} \quad (21)$$

where the observation x equals the time to lane crossing τ and $\ell_A(u; \tau)$ is the likelihood that CAAS action u will produce accurate (i.e., nominal or safe) consequences given τ .

2) *Liability*: From a strictly driver-autonomy perspective, whenever action interferes with driver autonomy then invoking a CAAS action is incorrect, and whenever action does not interfere with driver autonomy then invoking a CAAS action is correct. The liability function encodes the cost of incorrect CAAS action. Since all CAAS actions impact driver autonomy, any CAAS action has nonnegative driver cost.

The precise cost of the action depends on a) the probability that the action will produce fault consequences (effectual interference with driver autonomy) and b) the interference cost that accrues assuming that the action produces fault consequences. The probability that the action will interfere with driver autonomy is tantamount to the probability that the driver perceives a CAAS action as irrelevant to the current circumstances. Recall that a consequence is a function of action and driving state $c(u, \theta)$. The probability that action produces interference (fault) consequences $c(u, \theta) \in \mathbb{N}$ given the current state is given by $p(c(u, \theta) \in F | \theta) p_F(u | \theta)$ where F is the random variable

$$F = \begin{cases} u, & \text{if } c(u, \theta) \in F \\ 0, & \text{otherwise.} \end{cases}$$

The cost of an action is the increase in interference that results from taking action, and this is denoted by $J_L(u; \theta)$. Combining these factors gives

$$\mu_L(u, \theta) = J_L(u; \theta) p_F(u | \theta). \quad (22)$$

Taking the expected value of (22) with respect to $p(\theta | x)$ gives

$$\mu_L(u; x) = \sum_{\theta \in \Theta} J_L(u; \theta) p_F(u | \theta) p(\theta | x). \quad (23)$$

Within the scope of this paper, we assume that $p(\theta | x) = \delta(\theta - x)$ where δ is a delta function; this is equivalent to assuming that our observation x precisely determines θ . Given this assumption, (23) becomes

$$\begin{aligned} \mu_L(u; x) &= J_L(u; x) p_F(u | x) \\ &= J_L(u; \tau) \ell_L(u; \tau) \end{aligned} \quad (24)$$

where the observation x equals the time to lane crossing τ and $\ell_L(u; \tau)$ is the likelihood that CAAS action u will produce liable (i.e., fault or interfering) consequences given τ .

3) *Summary*: Equations (21) and (24) have the form identified in (25)

expected valuation = valuation \times likelihood

$$f(u; \tau) = J(u; \tau) \ell(u; \tau) \quad (25)$$

where *valuation* means either driver cost or safety payoff, respectively, and *likelihood* means the likely consequences given the measured time to lane crossing.

REFERENCES

- [1] R. W. Allen, "The driver's role in collision avoidance systems," in *Collision Avoidance Systems: Issues & Opportunities—Proc.*, Reston, VA, Mar. 1994, supported by ITS America Safety and Human Factors Committee and The National Traffic Safety Administration, pp. 33–51.
- [2] L. Bainbridge, "Ironies of automation," *Automatica*, vol. 19, no. 6, pp. 775–779, 1983.
- [3] M. Baker and M. Van Aerde, "Evaluation of transportation impacts arising from acc system deployment," in *Proc. 1997 Intelligent Transportation Systems Conf.*, Boston, MA, Nov. 1997.
- [4] E. R. Boer, E. C. Hildreth, and M. A. Goodrich, "Drivers in pursuit of perceptual and virtual targets," in *Proc. IEEE Intelligent Vehicles '98 Symp.*, Stuttgart, Germany, Oct. 28–30, 1998.
- [5] A. Burgett, "Crash avoidance holds key to safer US highways," *ITS: Intell. Transport. Syst.*, pp. 94–98, Sept. 1996.
- [6] L. Evans, *Traffic Safety and the Driver*. New York: Von Nostrand Reinhold, 1991.
- [7] J. Forbes, T. Huang, K. Kanazawa, and S. Russel, "The BATmobile: Toward a Bayesian automated taxi," *Int. Joint Conf. Artificial Intelligence*, submitted for publication.
- [8] L. M. Forbes, "DISCUSSION: The driver's role in collision avoidance systems," in *Collision Avoidance Systems: Issues & Opportunities—Proc.*, Reston, VA, Mar. 1994, sponsored by ITS America Safety and Human Factors Committee and The National Traffic Safety Administration, pp. 53–57.
- [9] G. Gigerenzer and D. G. Goldstein, "Reasoning the fast and frugal way: Models of bounded rationality," *Psychol. Rev.*, vol. 103, no. 4, pp. 650–669, 1996.
- [10] H. Godthelp, "Vehicle control during curve driving," *Human Factors*, vol. 28, no. 2, pp. 211–221, Apr. 1986.
- [11] H. Godthelp, P. Milgram, and G. Blaauw, "The development of a time-related measure to describe driving strategy," *Human Factors*, vol. 26, no. 3, pp. 257–268, June 1984.
- [12] M. A. Goodrich, "A theory of satisficing control," Ph.D. dissertation, Brigham Young University, Provo, UT, 1996.
- [13] M. A. Goodrich and E. R. Boer, "Maintaining driver autonomy with collision avoidance systems: A satisficing approach," Cambridge Basic Res., Nissan Res. Devel., Inc., Cambridge, MA 02142, Tech. Rep. TR-97-2, 1997.
- [14] —, "Multiple mental models, automation strategies, and intelligent vehicle systems," in *IEEE/IEE/JSAI Int. Conf. Intelligent Transportation Systems*, Tokyo, Japan, Oct. 5–8, 1999, pp. 859–864.
- [15] M. A. Goodrich, E. R. Boer, and H. Inoue, "Brake initiation and braking dynamics: A human-centered study of desired ACC characteristics," Cambridge Basic Res., Nissan Res. Develop., Inc., Cambridge, MA 02142, Tech. Rep. TR-98-5, 1998.
- [16] —, "A model of human brake initiation behavior with implications for ACC design," in *IEEE/IEE/JSAI Int. Conf. Intelligent Transportation Systems*, Tokyo, Japan, Oct. 5–8, 1999, pp. 86–91.
- [17] M. A. Goodrich, W. C. Stirling, and R. L. Frost, "A satisficing approach to intelligent control of nonlinear systems," in *Proc. 1996 IEEE Int. Symp. Intelligent Control*, Dearborn, MI, Sept. 15–18, 1996.
- [18] —, "A theory of satisficing decisions and control," *IEEE Trans. Syst., Man, Cybern.—Pt. A*, vol. 28, pp. 763–779, Nov. 1998.
- [19] —, "Model predictive satisficing fuzzy logic control," *IEEE Trans. Fuzzy Syst.*, vol. 7, pp. 319–332, June 1999.
- [20] Y. C. Ho, "Heuristics, rules of thumb, and the 80/20 proposition," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 1025–1027, May 1994.
- [21] —, "On the numerical solutions of stochastic optimization problem," *IEEE Trans. Automat. Contr.*, vol. 42, pp. 727–729, May 1997.
- [22] R. M. Hogan, "Impact of physical disengagement on driver alertness: Implications for precursors of a fully automated highway system," in *Proc. 1997 Intelligent Transportation Systems Conf.*, Boston, MA, Nov. 1997.
- [23] A. D. Horowitz and T. A. Dingus, "Warning signal design: A key human factors issue in an in-vehicle front-to-rear-end collision warning system," in *Proc. Human Factors Soc. 36th Annu. Meet.*, Atlanta, GA, Oct. 1992, pp. 1011–1013.
- [24] T. Inagaki, "Situation-adaptive responsibility allocation for human-centered automation," *Trans. Soc. Instrum. Contr. Eng.*, vol. 31, no. 3, pp. 292–298, 1995.
- [25] T. Inagaki and K. Inoue, "Adaptive choice of a safety management scheme upon an alarm under supervisory control of a large-complex system," *Reliab. Eng. Syst. Safety*, vol. 39, pp. 81–87, 1993.
- [26] T. Inagaki and M. Itoh, "Trust, autonomy, and authority in human-machine systems: Situation-adaptive coordination for systems safety," in *Proc. CSEPC 1996*, 1996, pp. 176–183.
- [27] —, "Trust, autonomy, and authority in human-machine systems: Situation-adaptive coordination for systems safety," *Inst. Inform. Sci. Electron.*, Univ. Tsukuba, Tsukuba 305, Japan, Tech. Rep. ISE-TR-96-138, Sept. 1996.
- [28] B. E. Kaufman, "A new theory of satisficing," *J. Behav. Econ.*, vol. 19, no. 1, pp. 35–51, 1990.
- [29] D. J. LeBlanc, G. E. Johnson, P. J. T. Venhovens, G. Gerber, R. DeSonia, R. D. Ervin, C.-F. Lin, A. G. Ulsoy, and T. E. Pilutti, "CAPC: A road-departure prevention system," *IEEE Contr. Syst. Mag.*, vol. 16, pp. 61–71, Dec. 1996.

- [30] D. J. LeBlanc, P. J. T. Venhovens, C.-F. Lin, T. E. Pilutti, R. D. Ervin, A. G. Ulsoy, C. MacAdam, and G. E. Johnson, "A warning and intervention system to prevent road-departure accidents," in *The Dynamics of Vehicles on Roads and Tracks, Proc. 14th IAVSD Symp.*, L. Segel, Ed., Ann Arbor, MI, Aug. 1996.
- [31] I. Levi, *The Enterprise of Knowledge*. Cambridge, MA: MIT Press, 1980.
- [32] G. Lilly, "Bounded rationality: A Simon-like explication," *J. Econ. Dyn. Contr.*, vol. 18, pp. 205–230, 1994.
- [33] T. Matsuda and S. Takatsu, "Algebraic properties of satisficing decision criterion," *Inform. Sci.*, vol. 17, no. 3, pp. 221–237, 1979.
- [34] —, "Characterization of satisficing decision criterion," *Inform. Sci.*, vol. 17, no. 2, pp. 131–151, 1979.
- [35] M. D. Mesarovic, "Systems theoretic approach to formal theory of problem solving," in *Theoretical Approaches to Non-Numerical Problem Solving*, R. Banerji and M. D. Mesarovic, Eds. Berlin, Germany: Springer, 1970, pp. 161–178.
- [36] M. D. Mesarovic and Y. Takahara, "On a qualitative theory of satisfactory control," *Inform. Sci.*, vol. 4, no. 4, pp. 291–313, Oct. 1972.
- [37] A. Meystel, *Autonomous Mobile Robots*. Singapore: World Scientific, 1991.
- [38] N. Moray, "Designing for transportation safety in the light of perception, attention, and mental models," *Ergonomics*, vol. 33, no. 10/11, pp. 1201–1213, 1990.
- [39] J. Murray and Y. Liu, "Hortatory operations in highway traffic management," *IEEE Trans. Syst., Man, Cybern.—Pt. A*, vol. 27, pp. 340–350, May 1997.
- [40] L. Nilsson, "Safety effects of adaptive cruise controls in critical traffic situations," in *Proc. Steps Forward, Vol. III*, Yokohama, Japan, Nov. 9–11, 1995, 2nd World Congr. Intelligent Transport Systems, pp. 1254–1259.
- [41] R. W. Proctor and T. Van Zandt, *Human Factors in Simple and Complex Systems*. Boston, MA: Allyn and Bacon, 1994.
- [42] J. Reason, "Cognitive aids in process environments: Prostheses or tools?," *Int. J. Man–Machine Studies*, vol. 27, pp. 463–470, 1987.
- [43] R. Rensink and E. Boer, Eds., "Cambridge Basic Research 1996 Annual Report," Nissan Res. Devel., Inc., Cambridge, MA, Tech. Rep. CBR TR 96-10, Dec. 1996.
- [44] S. Sen, Ed., "Satisficing Models," AAAI Spring Symp., Stanford, CA, Tech. Rep. SS-98-05, March 23–25, 1998.
- [45] H. A. Simon, "A behavioral model of rational choice," *Quart. J. Econ.*, vol. 59, pp. 99–118, 1955.
- [46] —, "Invariants of human behavior," *Annu. Rev. Psychol.*, vol. 41, pp. 1–19, 1990.
- [47] —, *The Sciences of the Artificial*, 3rd ed. Cambridge, MA: MIT Press, 1996.
- [48] M. Slotte, *Beyond Optimizing*. Cambridge, MA: Harvard Univ. Press, 1989.
- [49] W. C. Stirling, M. A. Goodrich, and R. L. Frost, "Procedurally rational decision-making and control," *IEEE Contr. Syst. Mag.*, vol. 16, pp. 66–75, Oct. 1996.
- [50] S. Takatsu, "Decomposition of satisficing decision problems," *Inform. Sci.*, vol. 22, no. 2, pp. 139–148, 1980.
- [51] —, "Latent satisficing decision criterion," *Inform. Sci.*, vol. 25, no. 2, pp. 145–152, 1981.
- [52] M. Tomizuka, J. K. Hedrick, and H. Pham, "Integrated maneuvering control for automated highway systems based on a magnetic reference/sensing system," Calif. PATH Res. Rep., Tech. Rep. UCB-ITS-PRR-95-12, 1995.
- [53] K. Vicente, "Should an interface always match the operator's mental model?," *CSERIAC Gateway*, vol. 8, no. 1, 1997. Published by Crew Syst. Ergonomics Inform. Anal. Ctr.
- [54] D. von Winterfeldt and W. Edwards, *Decision Analysis and Behavioral Research*. Cambridge, U.K.: Cambridge Univ. Press, 1986.
- [55] N. J. Ward, "Automation of task processes: An example of intelligent transportation systems," Paper based on ITS Focus Workshop entitled Intelligent Transport Systems and Safety, Feb. 29, 1996.
- [56] Y. Xiao, P. Milgram, and D. J. Doyle, "Planning behavior and its functional role in interactions with complex systems," *IEEE Trans. Syst., Man, Cybern.—Pt. A*, vol. 27, pp. 313–324, May 1997.
- [57] L. A. Zadeh, "What is optimal? (Editorial)," *IRE Trans. Inform. Theory*, vol. IT-4, p. 3, Mar. 1958.
- [58] —, "Maximizing sets and fuzzy Markoff algorithms," *IEEE Trans. Syst., Man, Cybern.—Pt. C*, vol. 28, pp. 9–15, Feb. 1998.
- [59] S. Zilberstein, "Using anytime algorithms in intelligent systems," *AI Mag.*, pp. 73–83, Fall 1996.
- [60] C. E. Zsombok and G. Klein, *Naturalistic Decision Making*. Hillsdale, NJ: Erlbaum, 1997.
- [61] Special Issue on Human Interaction with Complex Systems," *IEEE Trans. Syst., Man, Cybern.—Pt. A*, May 1997.



Michael A. Goodrich (S'92–M'97) received the B.S. (*cum laude*), M.S., and Ph.D. degrees in electrical and computer engineering from Brigham Young University, Provo, UT in 1992, 1995, and 1996, respectively.

From 1996 to 1998, he was a Research Associate with Nissan Cambridge Basic Research, Nissan Research and Development, Inc., Cambridge, MA, where he maintains status as a Visiting Scientist. Since 1998, he has been with the Computer Science Department at Brigham Young University where he is an Assistant Professor. His research interests include modeling and controlling intelligent systems, decision theory, multiple-agent learning and coordination, human-centered engineering, fuzzy logic, and estimation theory.



Erwin R. Boer (S'92–M'92) received the M.S. degree in electrical engineering from Twente University, The Netherlands in 1990, and the Ph.D. degree in electrical engineering from the University of Illinois, Chicago in 1995.

Since 1995, he has been a Research Scientist at Nissan Cambridge Basic Research, Cambridge, MA. His current research interests include human-machine interaction, control, and image processing and in particular their application in the development of human-centered operator-assist systems. Since 1990, he has been affiliated with the Center for Clouds Chemistry and Climate at Scripps Institution of Oceanography conducting research in the areas of remote sensing, image processing, and scientific data visualization.