Discretionary Behavior Switching: Analysis and Synthesis Results

Michael A. Goodrich and Thomas J. Palmer
Brigham Young University
Provo, Utah, 84602

Abstract

In previous work, we addressed how the world sometimes mandates switches in human behaviors [2]. This lead to a characterization of how humans manage such mandatory transitions. In addition to *mandatory* behavior switches, there are also situations where humans exhibit *discretionary* behavior switches. In this paper, we present a mathematical characterization of discretionary behavior switches which is applicable to both modeling human behavior generation as well as to developing action selection mechanisms for behavior-based robotics. We support this model by analyzing behaviors observed in human subjects and by synthesizing behaviors in a mobile robot.

1 Introduction and Previous Work

Sometimes changes in the environment mandate that people and robots change behaviors. However, even when the environment does not *mandate* a change in behavior, the environment can afford such a change and people/robots are free to choose to switch behaviors. In this paper, we present a mathematical framework for discretionary behavior switches. This framework uses the cost benefit tradeoff of satisficing decision theory, and includes the notion of goal-achieving/fault-avoiding behaviors. We take a somewhat unique approach to validating the model by not only *analyzing* data obtained from human subjects who participated in a driving experiment using the Nissan CBR driving simulator, but also *synthesizing* action selection mechanisms for a behavior-based robot architecture.

Multiple mental models. Many aspects of cognitive decision-making have been described in terms of mental models [4]. A mental model is an internal representation employed to encode, predict, and evaluate the consequences of perceived and intended changes to the operator's current state within the dynamic environment. We define a mental model $\mathcal{M}$ as a triplet consisting of the state of the environment Θ (including perceived and internal state components), a set of decisions or actions U, and a set of ordered consequences C that result from choosing $u \in U$ when $\theta \in \Theta$ obtains. According to this specification, a mental model not only encodes the relation between the input-action pair (θ, u) and the predicted consequence c, but also induces an evaluation of preferences among consequences.

Human cognition can be described using multiple mental models (treated as agents) which can be organized into a three level, multi-resolutional society of interacting agents corresponding to Rasmussen's knowledge-based (KB), rule-based (RB), and skill-based (SB) behaviors [6]. At the KB level of this hierarchy, the agent role is supervisory; at the RB level, the agent role is task management; and at the SB level, the agent role is task execution. Associated with each task-managing RB agent are a set of task-executing SB agents. Unlike RB agents which can be simultaneously enabled (multiple tasks can be simultaneously performed provided that attention can be suitably scheduled), only a single SB behavior can be enabled for each RB task. Since (a) only one SB behavior can be active and (b) no single SB behavior is sufficient for all conditions in the world, we require a mechanism for switching from one behavior to another.

Behavior-based Robotics. Employing bottom-up design, Brooks [1] proposed a layered architecture of increasingly sophisticated skills, and termed this architecture the subsumption architecture. Multiple skills are simultaneously active unless a higher level skill subsumes responsibility for the lower level behavior. In keeping with the spirit of behavior-based robotics, we assume the existence of a set of low level activities (like accelerating to a specified speed), but restrict attention to the management of skills (where skills are defined as patterned sequences of activities). These skills are managed by the task-management program. Rather than requiring the subsumption of one skill by a more sophisticated skill, we adopt a behavior-generation hierarchy similar to the human model discussed

0-7803-6583-6/00/$10.00 © 2000 IEEE

above and consistent with other approaches to behavior-based control (e.g., [5]). This allows us to treat robot action selection and switches between SB mental model agents in the same manner. We must therefore develop algorithms for switching between skills in a given a task context.

Satisficing decision theory. To develop algorithms for skill switching, we employ satisficing decision theory. The notion of satisficing was first identified by Simon [7] who 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*. Satisficing Decision Theory (SDT) is a multi-attribute extension [3] of Simon's approach which employs and compares two evaluation functions similar to the way benefit and cost are compared in economics literature. The key to this development lies in partitioning preferences over consequences into a generalized type of benefit called *acceptability*, and a generalized type of cost called *rejectability*.

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 acceptability $\mu_A : U \times \Theta \mapsto \Re$ and rejectability $\mu_R : U \times \Theta \mapsto \Re$ functions encode the preference relations defined for each consequence (i.e., action/state-of-nature pair). The SDT decision rule may be written as

$$S_b = \{(u;\theta) : \mu_A(u;\theta) \geq b\mu_R(u;\theta)\}, \qquad (1)$$

where b is a weighting parameter that represents the relative importance of cost and benefit. In SDT, preferences over consequences are represented by the benefit-like acceptability attribute and the cost-like rejectability attribute. These attributes are compared to determine when action u is admissible given state θ (i.e., when consequences are satisficing). For the speed management task in the context of automobile driving, the corresponding set of driver skills includes $U = \{\mathrm{CF}, \mathrm{SR}, \mathrm{AB}\}$, where CF indicates car following, SR indicates speed regulation (free driving), and AB indicates active braking. Also for the speed management task, the vector of perceptual states (see, for example, [2]) is $\theta = [T_c^{-1}, T_h, v_A]$ where T_c^{-1} is the inverse of time-to-contact, T_h is time headway, and v_A is the velocity of my vehicle.

Suppose that a human is performing skill $u \in U$. The human monitors θ via the RB mental model, and when $(u, \theta) \in S_b$ no change in skill-based behavior is necessary. However, when $(u, \theta) \notin S_b$, then the current behavior is not acceptable and must be switched to a behavior that is appropriate for the circumstances. The transition that occurs when θ produces satisficing consequences at time t but ceases to do so at time $t + 1$ is referred to as a perceptual triggering event. It is a perceptual occurrence which mandates a switch in behaviors, but the choice of which skill to choose is discretionary in nature.

2 Discretionary Switching

Skills must be changed when mandated by conditions in the world, but the environment does not always mandate which alternative skill should be selected. Additionally, humans do (and robots should) switch skills discretionarily even when the environment does not mandate a change.

2.1 Identification of Superior Alternatives

Satisficing, as we have defined it, is a notion of rationality determined by comparing two aspects of the consequences of making a decision. Under this rationality, a decision can be admitted or rejected without reference to other decisions. However, learning, memory, and the ability to model the world permits an agent to compare the consequences of one decision against another. This allows a decision maker to compare the consequences of alternative decisions in an effort to improve performance. In the two-attribute framework of satisficing decision theory, any option that is dominated by another option (i.e., the first option has lower acceptability and higher rejectability than the second option) can be eliminated. It is interesting to note (see [3]) that the set of non-dominated options is equivalent to the set of those options which maximize the hybrid utility $\alpha\mu_A(u;\theta) - (1-\alpha)\mu_R(u;\theta)$ for some tradeoff parameter $\alpha \in [0, 1]$. This means that the set of non-dominated options is equivalent to the set of maximizing options when the tradeoff parameter α is completely indeterminate. It also allows us to use simple maximization as the mechanism for discretionarily choosing a behavior once a decision to switch behaviors has been made.

2.2 Timing of Discretionary Switches

Whenever a perceptual triggering event occurs, we must perform a search for a satisficing alternative; for our purposes, this search is tantamount to finding the satisficing action that maximizes the difference between μ_A and μ_R. However, sometimes we search for superior alternatives even when our current behavior is satisficing. Such a search consumes attentional resources and should thus be undertaken only if it is likely to effect superior behaviors. We adopt the perspective that the likelihood of finding a superior behavior increases as the uninterrupted time spent executing a skill increases. We model this by initializing a random walk, and generating a search for dominating alternatives whenever the value of the random walk process exceeds a threshold.

3 Analysis: Automobile Driver Behavior Switching

In this section, we analyze the results of an experiment in which human subjects responded to cut-in events in a simulated automobile driving environment. We charaterize both mandatory and discretionary behavior switches.

3.1 Experiment Description

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 steering control but is constrained in longitudinal control. This constraint prevents the subject from driving more than ($v^* \approx 20\mathrm{m/s} \approx 43\mathrm{mph}$). Since subjects prefer to drive fast (and are encouraged to travel at maximum speed when safe), subjects usually travel at this maximum speed whenever permitted by traffic. Thus, subjects are nominally engaged in regulating speed unless the environment prevents them from doing so. During the experiment (and only when the driver is traveling at 20m/s), a cut-in vehicle passes the subject's vehicle and cuts into the lane with a ramdom initial time to contact $T_c-1(0)$ and randomly selected initial time headway $T_h(0)$. Data were partitioned into three classes corresponding to the three skills associated with the longitudinal vehicle control RB task: *active braking* AB (brake pedal depressed), *speed regulation* SR (car traveling at 20m/s), and *car following* CF (car accelerating or decelerating, but neither using brakes nor traveling at maximum speed).

3.2 Goal-Achievement versus Fault Avoidance

Skilled behaviors come in at least two qualitatively different varieties: goal-achieving and fault-avoiding[1]. A goal achieving-behavior is acceptable when it leads to the likely achievement of the goal and is rejectable when it leads to a likely fault state. A fault-avoiding behavior is acceptable when it leads away from a fault state and is rejectable when it leads to no known goal state. The fundamental goal of driving is usually to reach a destination, but this goal is constrained by the need to avoid risks. We will use this distinction to discriminate between acceptable and rejectable behaviors.

Within the context of the longitudinal vehicle control RB task, car following and speed regulation are goal-achieving skills; the long term goal is to reach a destination and this translates into a short term objective of traveling as fast as is safely allowed. Within the same context, braking is a fault-avoiding skill; the long term goal is to avoid injury and this translates into a short term objective of slowing down the vehicle to avoid a collision.

[1] "Exploring" seems to be a third qualitatively different behavior.

3.3 Dichotomous Values

The values μ_A and μ_R must be determined for each of the three skills SR, CF, and AB. We base these values on the overall goal of the driver. Therefore, both of the goal-achieving skills, SR and CF, are more acceptable at lower T_c^{-1}; it is desirable to move forward if conditions are safe. By contrast, AB is more acceptable in more dangerous circumstances (i.e., where T_c^{-1} is higher) because it acts to avoid or minimize a collision. Formally, μ_A is a cumulative measure of the percentage of times a particular skill is used along the T_c^{-1} axis, building towards the more acceptable direction. For each experimental trial, we extract the point at which each skill is initiated. The sets I_{SR}, I_{CF}, and I_{AB} signify the points where the skills SR, CF, and AB are initiated in the trials. The resulting acceptability values for each skill are:

$$\begin{aligned}
\mu_A(\mathrm{SR}, T_c^{-1}=\tau) &= \frac{N(T_c^{-1} \geq \tau | I_{\mathrm{SR}})}{N(T_c^{-1} \geq -\infty | I_{\mathrm{SR}})}, \\
\mu_A(\mathrm{CF}, T_c^{-1}=\tau) &= \frac{N(T_c^{-1} \geq \tau | I_{\mathrm{CF}})}{N(T_c^{-1} \geq -\infty | I_{\mathrm{CF}})}, \\
\mu_A(\mathrm{AB}, T_c^{-1}=\tau) &= \frac{N(T_c^{-1} \leq \tau | I_{\mathrm{AB}})}{N(T_c^{-1} \leq \infty | I_{\mathrm{AB}})},
\end{aligned}$$

where $N(T_c^{-1} \leq \tau | I_{\mathrm{skill}})$ is number of points in the set $\{\theta = [T_c^{-1}, T_h] \in I_{\mathrm{skill}} : T_c^{-1} \leq \tau\}$ (or conversely for $\geq$). These functions are shown in figure Figure 1.

Rejectability is based on where a particular skill ceases to be used. T_h, which reflects my car's proximity to the car in front of me, is used for calculating μ_R. As T_h decreases, the SR skill is terminated in favor of the AB or CF skill indicating that, in the presence of other traffic, speeding is likely to produce collisions; this means that rejectability of SR increases as T_h decreases. AB is similar, except that the skill is replaced by CF or SR as T_h increases. Again, this is reasonable as it is a fault avoiding behavior, but persisting in braking when no collision is imminent is counterproductive toward reaching the destination. Given

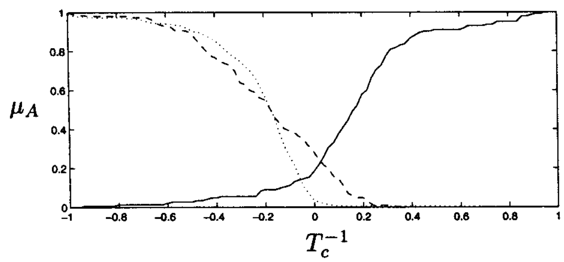

Figure 1: Acceptability: the solid line represents AB, the dashed line represents CF, and the dotted line represents SR.

the sets E_{SR} and E_{AB} as the sets of points where the SR and AB skills are terminated, respectively, the following functions specify μ_R:

$$\mu_R(\mathrm{SR}, T_h = \tau) = \frac{N(T_h \geq \tau | E_{\mathrm{SR}})}{N(T_h \geq -\infty | E_{\mathrm{SR}})},$$
$$\mu_R(\mathrm{AB}, T_h = \tau) = \frac{N(T_h \leq \tau | E_{\mathrm{AB}})}{N(T_h \leq \infty | E_{\mathrm{AB}})}.$$

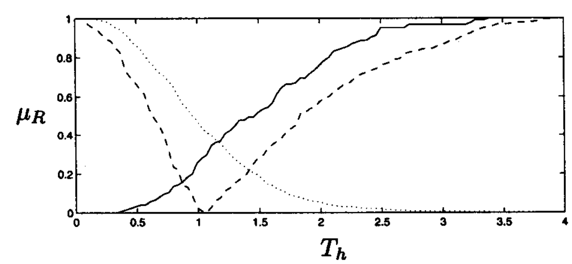

Figure 2: Rejectability: the solid line represents AB, the dashed line represents CF, and the dotted line represents SR.

CF is more complicated. In some cases, a driver terminates the CF skill in favor of AB and at other times, a driver terminates the CF skill in favor of SR. For each switch to AB, $\mu_R(\mathrm{CF})$ increases towards low T_h, and for each switch to SR, $\mu_R(\mathrm{CF})$ increases towards high T_h. Formally, let $E_{\mathrm{CF}}(\mathit{skill})$ denote the set of states for which CF is terminated and is replaced by *skill*. Because the experiment was not designed to test switches from CF to AB skills, we use the set I_{AB} (the set of points where braking is initiated) as an estimate of $E_{\mathrm{CF}}(\mathrm{AB})$. The rejectability of CF is given by

$$\mu_R(\mathrm{CF}, T_h = \tau) = \max\left[\frac{N(T_h \leq \tau | E_{\mathrm{CF}}(\mathrm{AB}))}{N(T_h \leq \infty | E_{\mathrm{CF}}(\mathrm{AB}))}, \frac{N(T_h \geq \tau | E_{\mathrm{CF}}(\mathrm{SR}))}{N(T_h \geq -\infty | E_{\mathrm{CF}}(\mathrm{SR}))}\right]$$

which results in the V-shaped function in Figure 2. The rejectability functions for the other skills are also shown in Figure 2.

3.4 Mandatory Switches

This accuracy of this classification scheme is given by the fraction of points in each starting set (the sets used to determine μ_A) that are considered satisficing. one free parameter: b. With $b = 1.0$, the number of points correctly classified as satisficing for SR, CF, and AB were 71%, 53%, and 85%, respectively. Clearly, we can find b (for example, $b = 0$) such that all points in I_{SR} and I_{CF} are satisficing, but this doesn't help us identify the boundary of the satisficing set. Instead, we want to match the satisficing boundary to the states where behaviors cease to be exhibited. We use a 75% classification accuracy as the standard for determining when we have closely matched the boundary. Decreasing b for SR to 0.96 and b for CF to 0.57, increases the classification accuracy to 75% for both SR and CF. Satisficing regions for these values of b are shown in Figure 3.

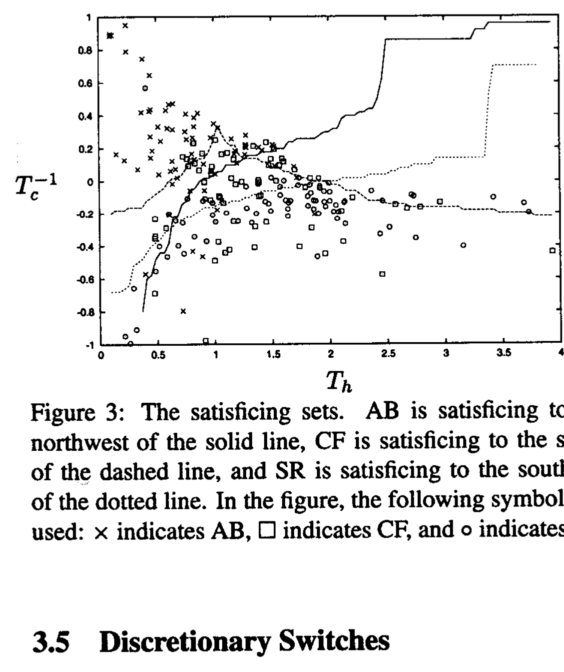

Figure 3: The satisficing sets. AB is satisficing to the northwest of the solid line, CF is satisficing to the south of the dashed line, and SR is satisficing to the southeast of the dotted line. In the figure, the following symbols are used: × indicates AB, □ indicates CF, and ∘ indicates SR.

3.5 Discretionary Switches

For the case of discretionary switching, a metric of superior alternatives is also needed. As stated earlier, nondomination is equivalent to optimization on the combined values of μ_A and μ_R. For the current data and with a tradeoff parameter of $\alpha = \frac{1}{2}$, this measure distinctly splits the state space into three areas where each skill dominates. The partitioning is shown in Figure 4, and represents a first order approximation to the regions where discretionary switches can take place. By comparing Figure 4 to Figure 3 we observe, for example, that although CF behaviors are satisficing in the region of $T_c^{-1} \approx -0.2$ and $T_h \approx 2$ and there is therefore no mandatory reason to

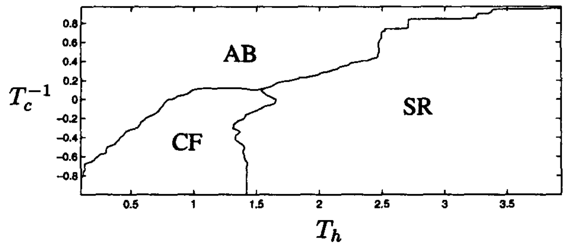

Figure 4: Approximate boundaries of the non-dominated sets. AB dominates in the northwest region, CF dominates in the southwest region, and SR dominates in the southeast region.

switch to SR, many SR skills are nevertheless present indicating a discretionary bias toward this skill. Although more work is needed to verify this trend, this indicates that non-domination has some correspondence with discretionary behavior switching.

4 Synthesis: Robot Behavior Switching

A robot operating in a multiple-robot society must be able to interact with other robots. Such interaction must include the ability to adapt robot speed in the presence of other robots (and humans) but still achieve a goal of reaching a destination. Based on the previous section, we decompose robot behaviors in a hierarchical manner, and focus on the longitudinal speed management task and three associated skill-based behaviors: robot following, speed regulation, and stopping.

4.1 Skills: Algorithm Descriptions

We implemented three robot behaviors that match the corresponding human skills: speeding, following, and stopping. *Speeding* was implemented simply by setting the velocity of the robot's wheels to 20 inches per second. If the velocity was less than this value, we increased the velocity by 2 inches per second each sample time until the desired velocity was obtained. *Following* was implemented using a PD controller. This controller operated on the difference between the actual time headway and a desired 1.5 second time headway. This skill produced a natural following pattern with the robot increasing following distance at high speeds, and decreasing it at low speeds. We limited the maximum acceleration to 1 inch per second each sample time, and limited the maximum deceleration to 5 inches per second each sample time. This maximum deceleration is insufficient to cause the robot to safely stop when it is travelling full speed and it encounters a stationary obstacle (like a wall); thus, the robot must switch to the stopping skill to avoid a collision. *Stopping* was implemented using the low level stop command. This command causes the robot's wheels to slow down as fast as possible without stressing the motors.

4.2 Experiment Description

Consider a hypothetical situation where a robot has identified a goal and is moving toward that goal. Suppose that the path to the goal is a straight line but that other robots must, at times, travel along this path too. Our objective is to have the robot travel to the goal efficiently (in minimum time) and safely (while avoiding collisions with other robots). In the experiment, the robot travels through a 30m hallway to a goal at the other end. At various random intervals, a second human-controlled robot cuts in front of the first robot's desired path. Using mandatory and discretionary behavior switching, the robot switches between its three skills.

4.3 Values

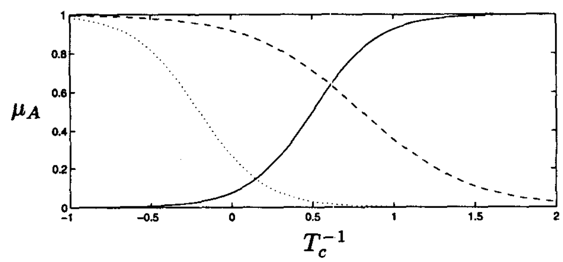

Figure 5: Acceptability: the solid line represents AB, the dashed line represents CF, and the dotted line represents SR.

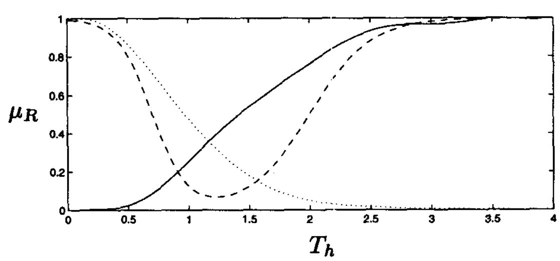

Figure 6: Rejectability: the solid line represents AB, the dashed line represents CF, and the dotted line represents SR.

Because the robot is much lighter than an automobile, we adapted the values to match the robot's dynamics. Plots of μ_A and μ_R for each robot skill are shown in Figures 5-6, respectively. These lead to the satisficing region shown in

Figures 7 and a non-dominated region qualitatively similar to Figure 4, but with a larger region of support for the *following* skill. For all skills, we used $b = 1$.

When a skill is no longer satisficing, the optimal skill is always switched to, since searching between three skills is not computationally intensive. Additionally, a random walk is used when a satisficing skill is in use to determine when discretionary switch is considered. It is interesting to note that even though the search time for an optimal skill is negligible for this experiment, it is not advisable to choose the "optimal" action at every time step. Data points on the boundaries between the optimal regions for two skills can result in undesirable switching (chatter) between the two skills.

4.4 Results

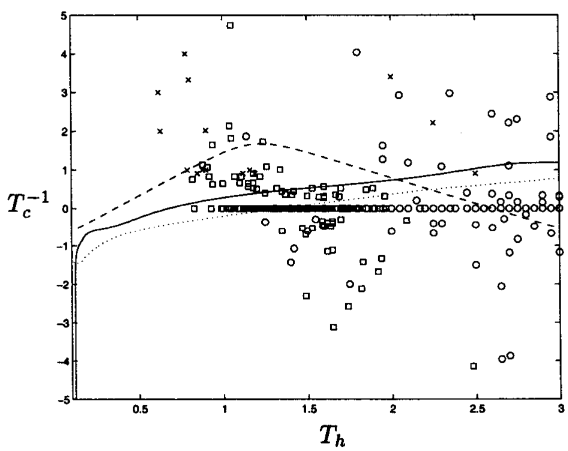

Figure 7: The satisficing sets. AB is satisficing to the northwest of the solid line, CF is satisficing to the southeast of the dashed line, and SR is satisficing to the southeast of the dotted line. In the figure, the following symbols are used: × indicates AB, □ indicates CF, and ○ indicates SR.

The μ_A and μ_R measures described above resulted in natural robot behavior, not too distinct from the human driving simulator results, as shown in Figure 7. showing the same satisficing regions shown above. Note that some skills were used outside of satisficing regions. This is because of the high degree of variance. Across the trials the mean T_c^{-1} was 0.118, while the standard deviation was 1.72. To compensate for such sudden noisy changes in T_c^{-1} and T_h, we place a delay on a switch to braking. If the environment mandates a change to AB, we record the state but do not perform the switch. If, on the subsequent time step, a switch to AB is still mandated, then the braking is performed. Hence, there are points for which braking appears to be required but is not actually used.

5 Summary

We presented a mathematical framework for discretionary behavior switches that was applied both to modeling human behavior generation and to robot action selection. We demonstrated how this framework could be used to explain and synthesize observed and desired behaviors.

References

[1] Rodney A. Brooks. A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, 2:14–23, 1986.

[2] M. A. Goodrich, E. R. Boer, and H. Inoue. A model of human brake initiation behavior with implications for ACC design. In *IEEE/IEEJ/JSAI International Conference on Intelligent Transportation Systems*, pages 86–91, Tokyo, Japan, October 5-8 1999.

[3] M. A. Goodrich, W. C. Stirling, and R. L. Frost. Model predictive satisficing fuzzy logic control. *IEEE Transactions on Fuzzy Systems*, 7(3), June 1999.

[4] P. N. Johnson-Laird. *The Computer and the Mind: An Introduction to Cognitive Science*. Harvard University Press, Cambridge, Massachusetts, 1988.

[5] P. Pirjanian and M. Matarić. A decision-theoretic approach to fuzzy behavior coordination. In *IEEE International Symposium on Computational Intelligence in Robotics and Automation*, November 1999.

[6] J. Rasmussen. Outlines of a hybrid model of the process plant operator. In T. B. Sheridan and G. Johannsen, editors, *Monitoring Behavior and Supervisory Control*, pages 371–383. Plenum, 1976.

[7] H. A. Simon. A behavioral model of rational choice. *Quart. J. Economics*, 59:99–118, 1955.