Automotive Research Center
A U.S. Army Center of Excellence for Modeling and Simulation of Ground Vehicles
Led by the University of Michigan

Modeling Bi-Directional Trust in Semi-Autonomy for Improved System Performance
Quad Members: Dawn Tilbury, Lionel Robert, Xi Jessie Yang; Hebert Azevedo-Sa, Suresh Kumaar Jayaraman, Connor Esterwood (University of Michigan)
Victor Paul, Ben Haynes, Dariusz Mikulski (CCDC/GVSC); Mitch Rohde (Quantum Signal, LLC)

Motivation

- Automated Vehicles (AVs) can improve drivers' safety and non-driving related task performances, but trust is necessary
- However, trust miscalibrations (mismatches between drivers’ trust in the AV and AV’s capabilities) lead to inappropriate use of the AV, as
  - Undertrust induces *disuse* of the AV
  - Overtrust induces *misuse* of the AV
- Army-GVSC relevance: trust needs to be managed in soldier-autonomy teams

Fundamental Research Questions

- Can we describe how drivers' trust in the AV evolves over time?
- Can we design a system that interacts with drivers to reduce trust miscalibrations?

Objectives

- Model the dynamics of drivers' trust in the AV, identifying
  - factors that influence drivers' trust in the AV
  - how trust varies over time, while the AV interacts with the driver
- Develop a framework to manage drivers' trust in the AV, consisting of
  - trust estimation (from drivers' behavioral cues and actions)
  - trust control (manipulating drivers' trust to reach trust calibration)

Approach

- Human-in-the-loop experiments with driving simulator
- Visual search non-driving related task to assess drivers' performances and determine cash bonuses
- Eye-tracking to measure drivers focus on the non-driving related task and on the driving task

Results & Discussion

1. Trust Model

   ![Trust Model Diagram](image)

   \[T(t+1) = AT(t) + Bz(t) + \frac{F(t)}{M(t)} + \omega\]

   Alarms from the AV

   Drivers' Trust in the AV

   Drivers' behavioral cues and actions

   Situation awareness and risk perception are crucial in trust development and trust dynamics model (considering the factors that influence trust) [1, 2]

2. Trust Estimation

   Successful trust estimation from drivers' behavioral cues and actions [3]

3. Trust Control

   Framework to increase trust of undertrusting drivers and decrease trust of overtrusting drivers, to avoid trust miscalibrations [4]

   ![Trust Control Diagram](image)

<table>
<thead>
<tr>
<th>Driver's trust state before interaction</th>
<th>Average trust variation after interaction (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undertrusting</td>
<td>+ 17.1 (2.7)</td>
</tr>
<tr>
<td>Calibrated trust</td>
<td>+ 1.7 (3.3)</td>
</tr>
<tr>
<td>Overtrusting</td>
<td>- 7.3 (3.5)</td>
</tr>
<tr>
<td>Extremely overtrusting</td>
<td>- 21.2 (5.5)</td>
</tr>
</tbody>
</table>

   Obs.: all parameters significant (p < 0.001)

Conclusion

- AVs can interact with drivers to calibrate drivers' trust and try to avoid disuse or misuse of automated driving, improving the human-autonomy team's safety and performance

References:


Fig. 1 - Experimental Setup

Simulated Driving Task
Non-Driving Related Task
Eye Tracker

Fig. 2 – Trust modeling

Fig. 3 – Trust Estimation Results

Table I – Trust Management Results