Ensuring the Safety of On-Road Self-Driving Car Testing

Prof. Philip Koopman

Carnegie Mellon University

PA AV Summit
April 9, 2018

Tempe Arizona / March 18, 2018

Carnegie Mellon University

Elaine Herzberg
Pre-impact dashcam image
Tempe Police Dept.

- **A tragic death has occurred**
  - How can we mitigate risk in the future?
- **Activities that do NOT improve safety of autonomous vehicle (AV) testing:**
  - Assigning blame
  - Arguing that delaying deployment costs lives
  - Finding out why autonomy failed (surprise!)

- **We should NOT sacrifice at-risk population for sake of progress**
  - **Instead, make progress with safe AV testing platforms**
    - AV testing platform = autonomy + safety driver + safety support technology

How Do You Know It's Safe Enough?

Carnegie Mellon University

- **Safety Case:**
  A structured **written argument**,
  supported by **evidence**,
  justifying system is **acceptably safe**
  for intended use.

National Transportation Safety Board/Handout via REUTERS

- **Example structure:**
  - Safety Reason 1 / evidence for reason 1
  - Safety Reason 2 / evidence for reason 2
  - Safety Reason 3 / evidence for reason 3
  - …

Safety Case Elements for AV Testing

Carnegie Mellon University

- **Essential observations for AV testing**
  - We care about safety of test vehicle
    - Autonomy is immature –
      that's why there is a safety driver!
  - Appropriately safe does not mean perfect

https://goo.gl/YUC5oU

- **AV testing safety goal: *no worse than human-driven vehicle***
  1. The safety driver is paying adequate attention
  2. The safety driver has time to react if needed
  3. When the safety driver reacts, the vehicle will respond properly

Is the Safety Driver Really In the Loop?

Carnegie Mellon University

- **"We have a safety driver" doesn't cut it as an argument**
- **Driver Dropout is well known**
  - Airline pilots (even if there are two!)
  - 1990s-era Automated Highway System
  - Can't just assume alert safety drivers
- **Questions to ask about safety drivers:**
  - Are they trained?
  - How will you ensure they are alert/awake?
  - How will you monitor on-road performance?

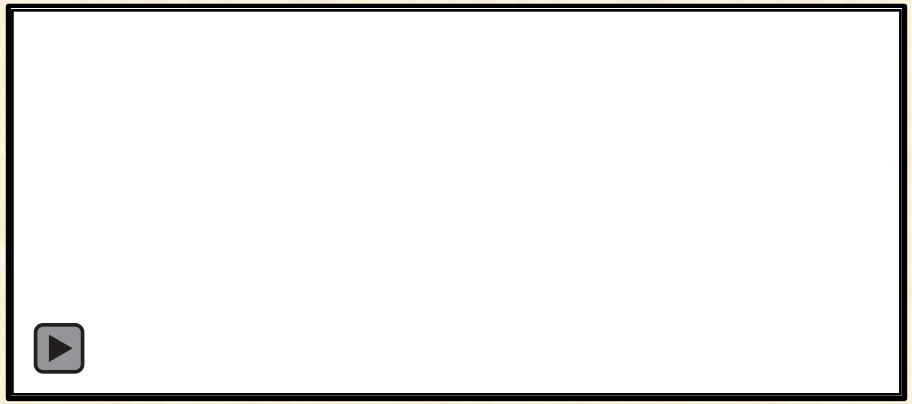
Snooze cruise: How the drama unfolded as the two pilots 'slumbered at the controls' 2009 https://goo.gl/5htvnP

Fernando De la Torre CMU: https://goo.gl/bHDXLS

Can Safety Driver React In Time?

Carnegie Mellon University

- **Safety Driver Tasks:**
  - Mental model of “normal” AV
  - Detect abnormal AV behavior
  - React & recover if needed



Jan 20, 2016; Handan, China

- **Example: obstructed lane**
  - Does driver know when to take over?
  - Can driver brake in time?
    - Or is sudden lane change necessary?

- **Example: two-way traffic**
  - What if AV commands sudden left turn into traffic?

Keeping the Safety Driver in the Loop

Carnegie Mellon University

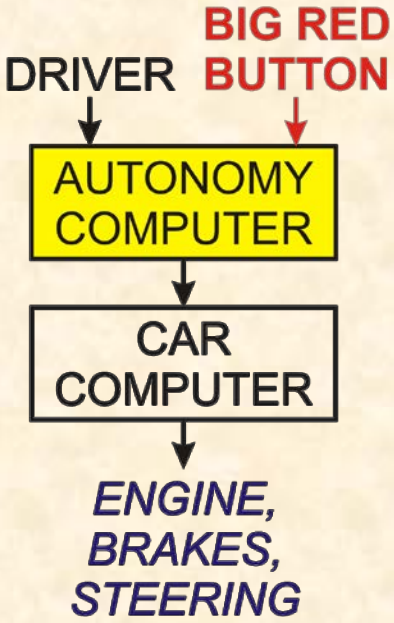
- **Supervisory human process:**
  - *First* detect AV problem; *then* react
- **Driver awareness of AV state**
  - Does AV see a pedestrian?
  - Is AV planning to avoid obstacle?
  - Is AV accurately displaying its intended plan?
- **Driver situational awareness**
  - Must intervene before it's too late to recover

Edge Case Research

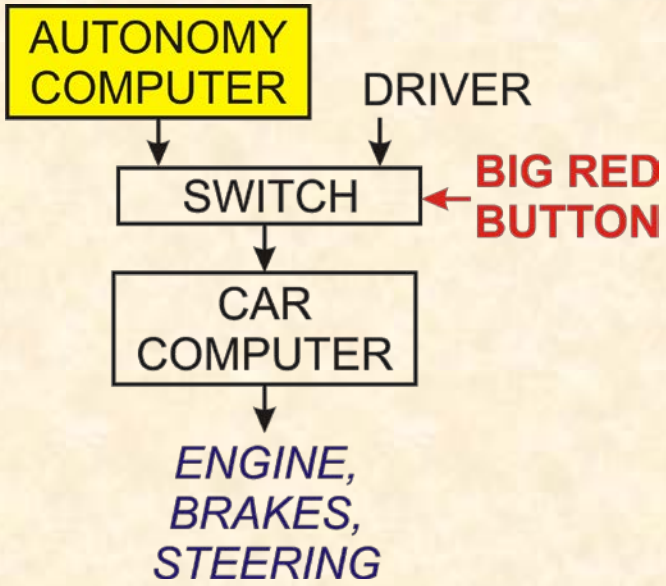
Does The Big Red Button Work?

Carnegie Mellon University

- **Claim: safety driver can over-ride autonomy**
- **Is this safe?**



Is this?



- **Use accepted practices to ensure disengagement safety**
  - For example, safety standard (ISO 26262) for disengagement mechanism

Example Safety Argument Sketch

Carnegie Mellon University

- **Safety driver(s) attentive**
  - Safety driver training, qualification
  - Real-time driver alertness monitoring
  - Review of driver performance data
- **Effective safety driver reaction**
  - Leave margin for recovery
  - Don't paint human driver into a corner
- **AV disengagement mechanism really works**
  - Follows safety engineering practices

Implementation Considerations

Carnegie Mellon University

- **Minimal regulatory intervention approach:**
  - AV testers provide the safety argument
    - Measured against criteria they themselves create
  - Who decides sufficiency?
    - Perhaps public review and litigation exposure

- **Key features of this safety approach:**
  - Proprietary autonomy information not revealed
  - Designer flexibility in choosing approach
  - Emphasizes adequate testing safety, not AV perfection

https://goo.gl/YUC5oU

Summary

Carnegie Mellon University

- **Proposed Safety Goal:**
  AV testing as safe as a human-driven vehicle
  1. Show that the safety driver is paying adequate attention
  2. Show that the safety driver has time to react if needed
  3. Show that AV disengagement/safing actually works

QUESTIONS?