Cliff Notes for Session 1 Presentations

- Missy Cummings
  - Experimental research involving 4 tests, 3 Tesla Model 3s vehicles on [name test track]
  - **Question**: How well do Level 2+ vehicles alert distracted drivers under various conditions?
  - Results run counter to Tesla's claim that running on autopilot is safer than not doing so
- Marjory Blumenthal
  - Issues for assessing and communicating about AV safety - Level 4 focus
  - Builds on 2018 measure framework: https://www.rand.org/pubs/research_reports/RR2662.html
  - Three principal approaches – Measurement, processes, thresholds (quantitative or qualitative)
  - Communicating about AV safety
- Ben Shneiderman
  - Aim: To "reframe thinking" with regard to human-machine interaction
  - Human-centered AI - 6 ingredients
  - 2D HCAI framework - RST systems require highly-human, highly automated control
  - Governance structures for 2D HCAI

Discussion Period

- Many good questions and comments wrt all 3 presentations
  - Directly to a speaker by a participant
  - Entered in Chat
- As with discussions at our in-person workshops
  - Many more questions/comments than the allotted time permits
  - Indeed, in overtime, several questions were asked a speaker after he/she had left the meeting
  - Please use email or other means to follow-up with new or unanswered questions of a speaker
- Thank you, Session 1 Chair Kevin Driscoll!
  - Very smooth and seamless management of both speaker intros and the subsequent discussion

Relevance to Workshop Goal

- Debate and provide arguments on all sides of the following hypothesis:
  - *L3 vehicles cannot be made acceptably safe with current technology and practices*
- Missy - Some experimental results that support “cannot” (for L2+)
- Marjory - Measure thresholds aimed at quantifying “acceptably safe”
- Ben - L3 vehicles not in high human, high auto RST quadrant - support
- My two cents
  - Regarding in-vehicle control of a road vehicle, human vs. autonomous control is 1-dimensional
  - In case there’s a proper mix of the two, as in L3
    - Successful handovers from one to the other become problematic
    - In turn, acceptably safe operation can be compromised
  - Experience with aircraft flight control systems is similar
    - Many accidents where mixed mode operation was to blame
    - Most recent example - MCAS problem with B737 MAX

Autonomous Vehicle Industry Perspectives

- **Technical Safety Challenges**
  - Safety assurance vs. certification
  - Autonomy levels vs. V&V and certification costs
  - V&V of AI/ML based functions
    - Perception, object detection, path planning, and prediction

- **Approaches to Quantifiable and Acceptable Safety:**
  - Safety Performance Indicators (SPI) for quantitative safety claims
  - Testability, dual redundancy, HW/SW/sensor diversity, high availability at mission critical times
  - Safety watchdogs, safety kernels, safety co-pilots
    - Independent simplified invariant checkers (e.g., collision, instability, lane departure, speed limit)
  - Reaction and recovery
    - Fail-safe mechanisms, graceful degradation, raise alerts, pull over to road-side

=> L3+ AD: An evolutionary process starting with success in simpler operational design domains, requiring new standards, new technologies for V&V and certification, and coopetition between industry and academia.

Rapporteur’s notes on 1/31/21 IVDS session

How do we know / can we assure that an AV is safe?

  - Lorenzo: with formal statistical methods incorporating conservative Bayesian inference (CBI) and ”bootstrapping” confidence based on operation without mishaps
  - Sanjit: with simulation-based falsification, scenario simulation in combination with verification
  - John: by building systems that employ generative modeling of the world and use them to detect surprise and respond

Comment: deployment of AVs at present seems to be made tolerable by limiting the operational environment. Not sure these talks really addressed this aspect as much as they might; perhaps an area for further refinement of models.

- What do they have to say about the workshop hypothesis: Resolved: ***Level 3 autonomous vehicles cannot be made acceptably safe with current technology and practices.***
  - It seems that a successful L3 system has to detect when it needs to handoff control to the driver
  - This seems like detection of surprise
  - So does the ability to build a proper L3 system actually imply we can build an L5 systems? Or perhaps we can’t build an L3 system until we already built an L5 system?

https://ivds2021.dependability.org

WG.10.4

First IFIP Workshop on Intelligent Vehicle Dependability & Security: Path Forward

Jan 31, 2021

Workshop Chair

Dr. Jay Lala

Sr. Principal Engineering Fellow

Raytheon Technologies

San Diego, CA

Stakeholder Engagement

- Industry: technology suppliers to automotive industry
- Academia
- Non-profits & government consultants
- Research institutions
- Standards influencers/shapers
- Automotive industry
- Regulatory and governance bodies

Good start but we need to do better

Technology & Societal Issues

- We know how to built fault-tolerant systems
  - We have been doing it for 40+ years for many different domains
  - Affordable fail-safe (not fail-stop) autonomous vehicle control systems are technically feasible
- Challenge is the adoption and implementation by automotive industry
  - Needs government & regulatory push, consumer pull and nudging by all other stakeholders
  - Last resort: accidents and lawsuits
- Application of Machine Learning algorithms still has ways to go before being deployed in safety-critical systems
- We don't know how to build cyber-resilient / intrusion-tolerant systems to the same degree

Fault-Tolerant designs a good foundation but many new challenges to overcome

Thank You!!!

- Speakers, Panelists and Session Chairs
- WG10.4 Friends and Guests
- My colleagues on the Organizing Committee: John, Carl, Chuck, and Homa