SAE INTERNATIONAL

WCX WORLD CONGRESS EXPERIENCE™

APRIL 10-12, 2018 • COBO CENTER • DETROIT, MICHIGAN

sae.org/wcx

Toward a Framework for Highly Automated Vehicle Safety Validation

Dr. Philip Koopman & Mike Wagner

Carnegie Mellon University

2018-01-1071 / 18AE-0273

The First 90% Is The Easy Part

WCX WORLD CONGRESS EXPERIENCE

But, the second 90% is the hard part.

1. **Be smarter than a billion miles of testing**
2. **Beware of simulation fidelity nirvana**
3. **Be sure tests pass for the right reason**
4. **Explicitly manage uncertainty**

https://goo.gl/oYnzY3

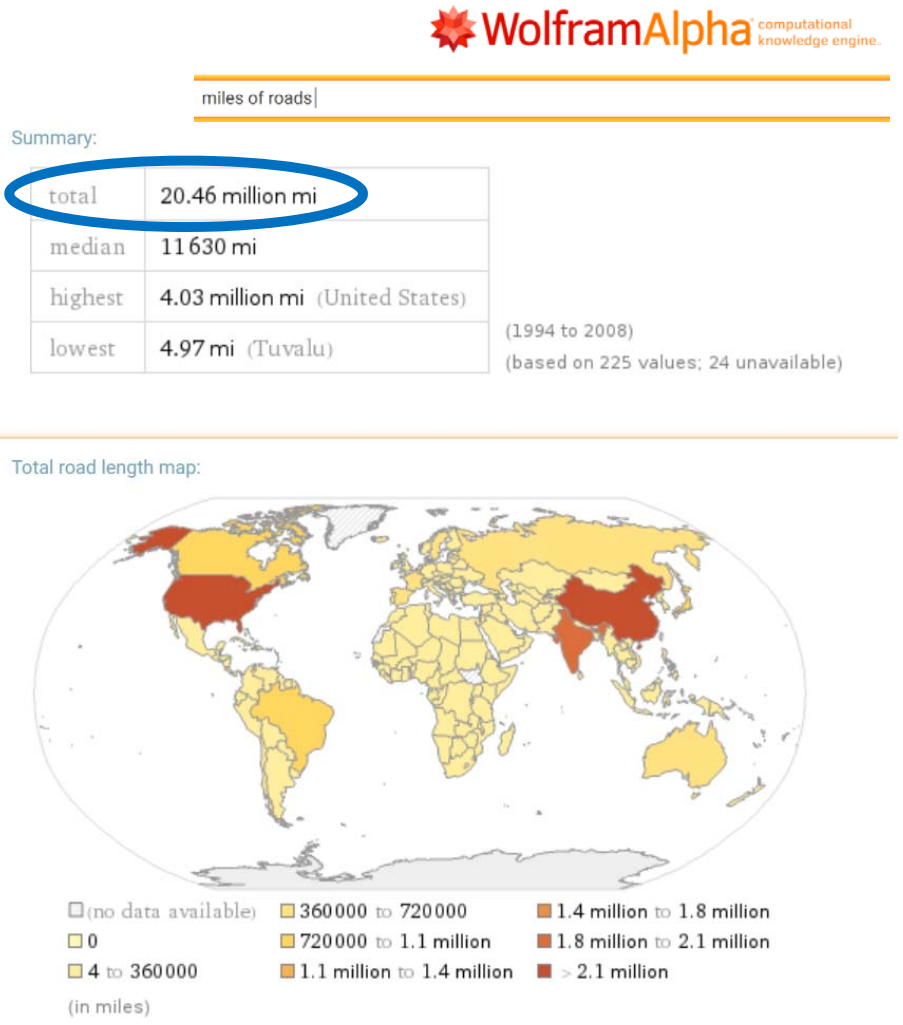
Do We Need Billions of Test Miles?

- **If 100M miles/critical mishap…**
  - Test 3x–10x longer than mishap rate
    ➔ Need 1 Billion miles of testing

- **That's ~25 round trips on every road in the world**
  - With fewer than 10 critical mishaps

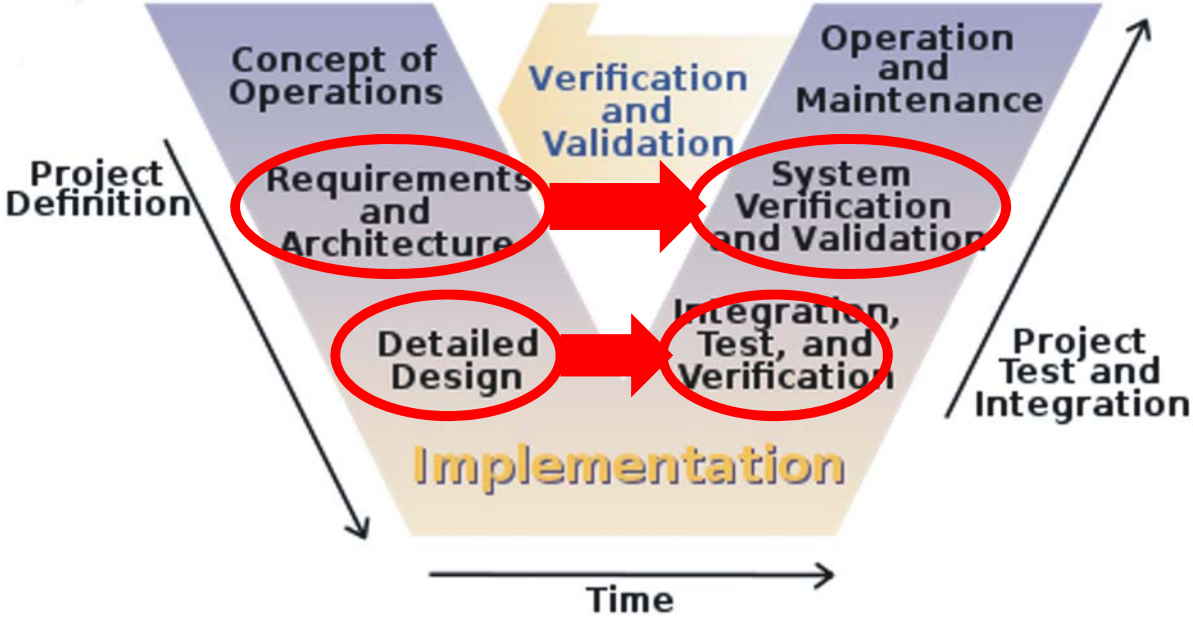
  …

  - *Then* you're only as good as a human
    - (Including the impaired humans!)



Traditional Validation Doesn't Need 1Gmi

- **If you have requirements and understand design:**
  - ISO 26262 for safety functions
  - Emerging SOTIF standards
- **Testing looks for holes in engineering rigor**
  - ***You should do this for everything you can!***



What If Traditional V Doesn't Seem To Fit?

- **Machine Learning (inductive training)**
  - No requirements
    - Training data is difficult to validate
  - No design insight
    - Generally inscrutable
    - Prone to over-fitting/gaming

https://en.wikipedia.org/wiki/Magic_Roundabout_(Swindon)

https://goo.gl/3dzguf

- **Use your road miles to gather requirements**
  - Novel objects, events, scenarios (OEDR-centric)
  - Novel operating conditions (ODD-centric)
  - Edge cases that present problems
  - Look for novelty even if your vehicle "test" is passing

- **Think "requirements testing" not "vehicle testing"**
  - Disengagements are a blunt instrument for detecting novelty

https://goo.gl/J3SSyu

Smart Use of Simulation

- **Point of view: everything is a simulation**
  - Software component simulation
  - Software vehicle simulation
  - HIL testbeds
  - Closed course testing
    - Simulated environment, obstacles, events
  - Public road testing
    - Assumes representativeness

University of Michigan

- **Even a "perfect" simulation needs scenarios as inputs**
  - You need a test plan that covers all required functionality

All Simulations Are "Wrong"

But some simulations are useful

- **It's all about the assumptions**
  - "Perfect" simulation is expensive
  - Exploit the cost/fidelity tradeoff
- **Layered Strategy:**
  - Simplified simulations explore large spaces
  - Complex simulations address residual risks
    - Validate assumptions made by simple simulations
    - Look for emergent effects and surprises
- **Use road tests to validate simulations**
  - Identify and concentrate simulation residual risks

Validation Activity	Residual Risks (Threats to Validity)
Pre-deployment road tests	Unexpected scenarios, environment
Closed course testing	*As above, plus:* Unexpected human driver behavior, degraded infrastructure, road hazards
Full vehicle & environment simulation	*As above, plus:* simulation inaccuracies, simulation simplifications (e.g., road friction, sensor noise, actuator noise)
Simplified vehicle & environment simulation	*As above, plus:* inaccurate vehicle dynamics, simplified sensor data quality (texture, reflection, shadows), simplified actuator effects (control loop time constants)
Subsystem simulation	*As above, plus:* subsystem interactions

Table 1. Hypothetical validation activities and threats to validity.

How Do You Know a Test Passed?

- **Traditional test paradigm:**
  - You think design is right
  - Test validates engineering done properly
    - Test traces to requirements/design
    - Deterministic behavior according to test plan

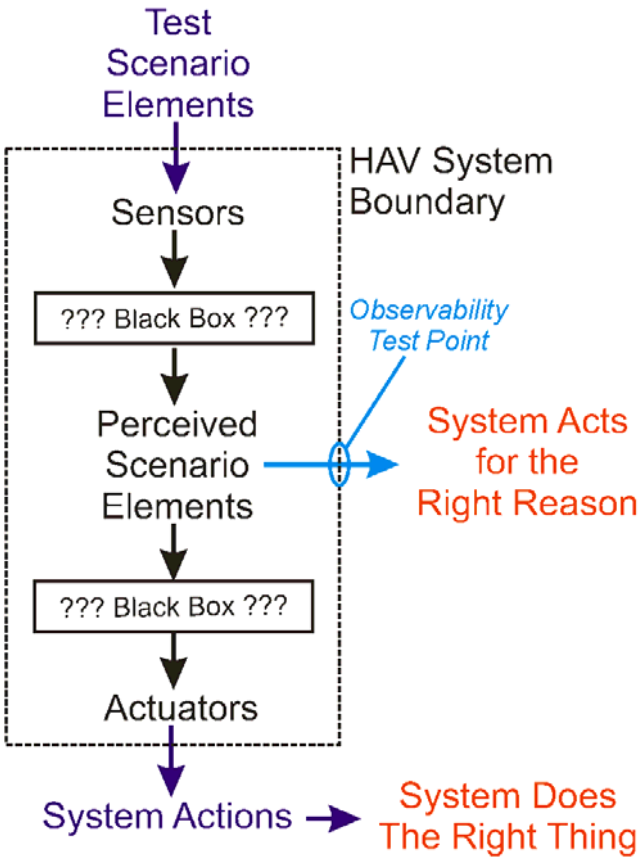
- **Inductive training test paradigm:**
  - You think system was trained properly
  - Test determines whether training worked
    - Weak traceability to test set, if any
    - Hope to detect training data gaps, overfitting
  - ***BUT:*** nondeterministic, opaque "design"

https://goo.gl/cFCknY

https://goo.gl/QdTYVV

Improving Observability for Testing

- **Hypothetical test:**
  - 10 tests of child in crosswalk
    - 10 times vehicle does not hit child
    - Conclusion: vehicle does not hit child in crosswalk
  - Threats to validity
    - Random path planner got lucky 10 times in a row
    - Vehicle only recognizes children in certain conditions
    - Vehicle thought a bush at that intersection is a child
    - …
- **Increase confidence via self-reporting**
  - Vehicle self-reports: "I see a child in a crosswalk"
    - Perception simulation: children, crosswalks, fuzzing
    - Vehicle simulation: simulated children/crosswalks
    - Test track: simulated children; real crosswalks
    - On-road testing: real children/crosswalks (with safety supervision!)



Explicitly Manage Uncertainty

- **Things we don't think matter**
  - But we might be wrong
- **Things we think are rare**
  - e.g., lightning strikes
    - But we might be wrong about that!

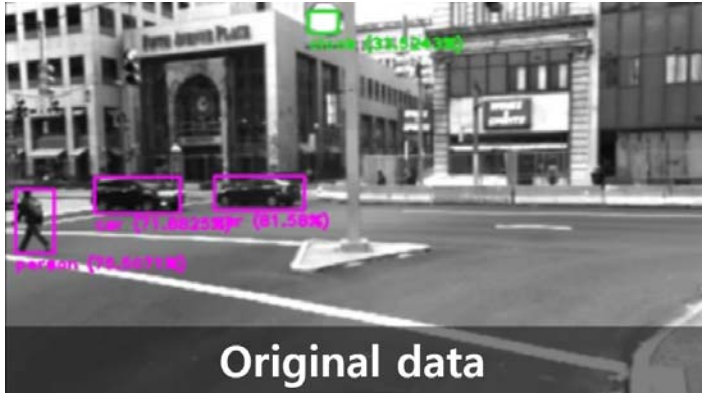
YouTube: PknOqXqcnUo, M1XHjl_6HtM, -0hE6gAcbvg, y6Krr4TazMg

- **Things we aren't completely sure about**
  - e.g., frequency of correlated sensor failures
  - Monitor quality of estimates
- **Things we didn't think of**
  - Try to detect "vehicle is clueless" (it's an ODD violation)
  - Do something reasonably safe

https://goo.gl/MZWGi1

Techniques for Managing Uncertainty

- **Do aggressive fault injection**
  - Even "unrealistic" faults provide insight
  - Especially important is perception fuzzing
    - Perturb both ODD and OEDR aspects of sensors



- **Document and monitor your assumptions**
  - "X" won't happen – put in a detector for "X"
  - "Y" is rare – measure arrival rate of "Y"
  - System will never do "Z" – test via fault injection

Pedestrian Missed: Gaussian Noise + Black Car

- **"We thought of everything"**
  - No. You didn't.

Pedestrian Missed: Gaussian Blur

Making the Second 90% Easier

WCX WORLD CONGRESS EXPERIENCE

1. Concentrate on data collection with road miles

- Look for things beyond disengagement triggers
- Use vehicle “testing” to validate simulations

2. Use a layered approach to simulation

- Exploit fidelity/cost tradeoffs
- Validate assumptions & simplifications

3. Monitor tests passing for the right reason

- Have system self-report scenario it thinks it is in

4. Monitor assumptions and surprises

- Actively look for having missed something