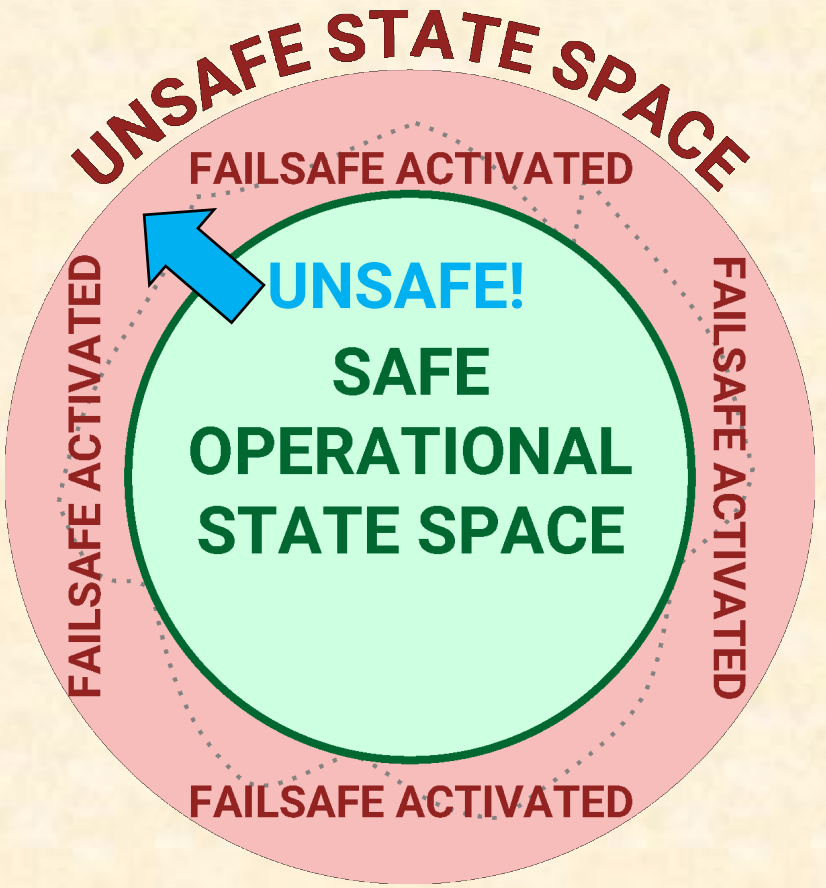
# Safety Architectures

# Safety Envelope Approach to ML Deployment



- Specify unsafe regions
- Specify safe regions
  - Under-approximate to simplify
- Trigger system safety response upon transition to unsafe region
- Inherent tension of envelope simplicity vs. permissiveness



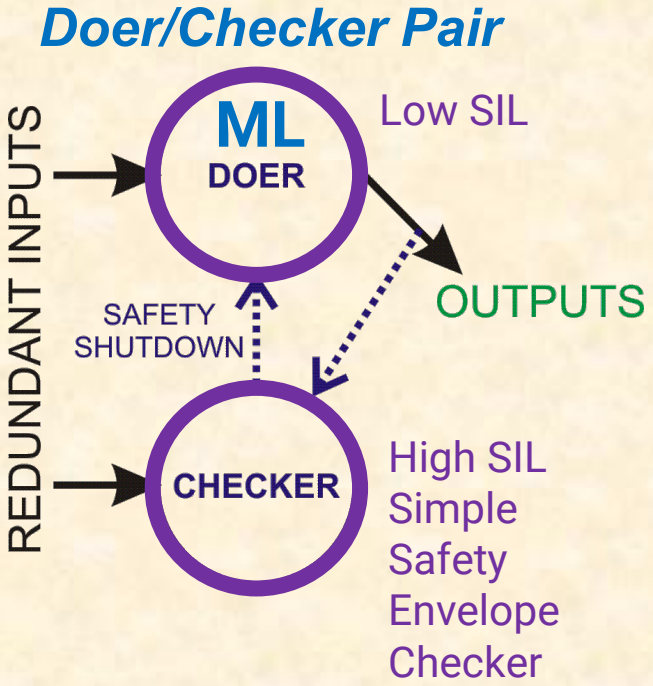

 

# Architecting A Safety Envelope System



- "Doer" subsystem
  - Implements normal, untrusted functionality
- "Checker" subsystem – Traditional SW
  - Implements failsafes (safety functions)
- Checker entirely responsible for safety
  - Doer can be at low Safety Integrity Level
  - Checker must be at higher SIL

(Also known as a "safety bag" approach or monitor/actuator pair)



# Self-driving shuttle company ordered to stop carrying passengers after injury

9

*The DOT suspends a shuttle operator on the same day it was criticized for being too hands-off*

By Sean O'Kane | @sokane1 | Feb 26, 2020, 12:56pm EST

SHARE

https://bit.ly/3fCdlpO


Photo by Paco Freire / SOPA Images/SOPA Images/LightRocket via Getty Images

# Self-driving shuttle company adds seatbelts in order to resume US operations

6

*But passenger rides might be scarce during the pandemic*

By Sean O'Kane | @sokane1 | May 18, 2020, 4:31pm EDT

SHARE

https://bit.ly/3ucNDTI


## SELF-DRIVING SHUTTLE BUS IN SPAIN'S MADRID PROVOKES CRASH ON FIRST DAY

*The vehicle was travelling at a speed of 20 kilometres per hour through the Universidad Autonoma de Madrid when it provoked the accident.*

By Cristina Hodgson - 23 Oct, 2020 @ 10:00 0

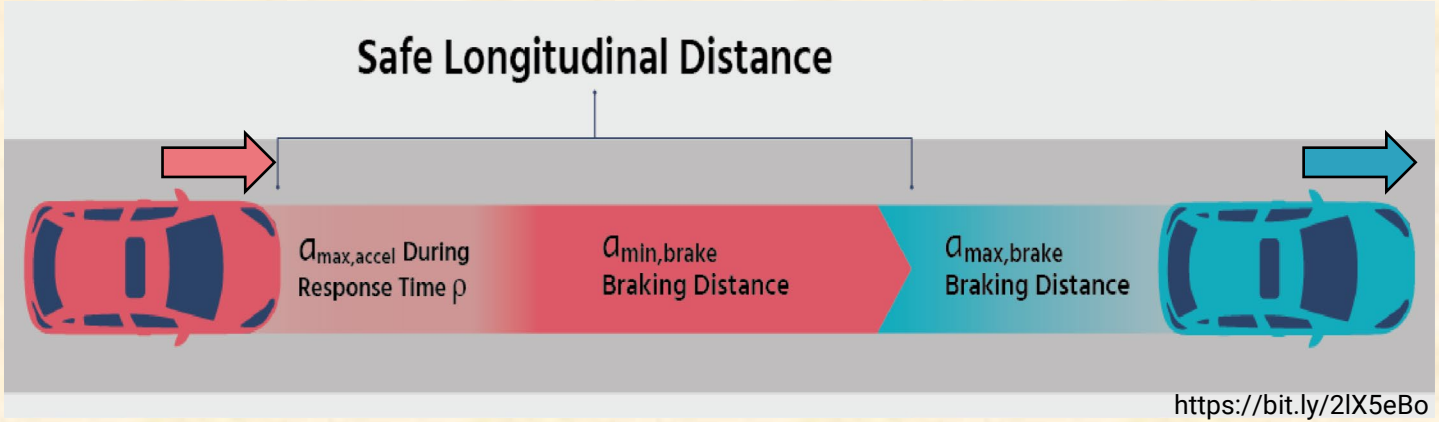

https://bit.ly/3udw3ie

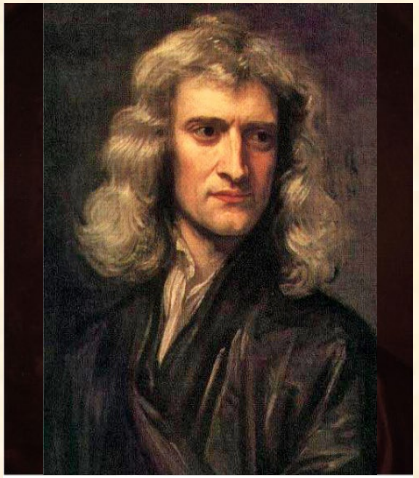
# Physics-Based Checker Rules



- Responsibility-Sensitive Safety (RSS) :
  - Safe distances based on physics
  - Defines proper responses to imminent collision




https://bit.ly/2lX5eBo



**F=MA**

It's not just a good idea.

**It's the Law!**

- Proofs don't eliminate uncertainty
  - Need knowledge of environment & other vehicle equipment capabilities

# Uncertainty in the World Model

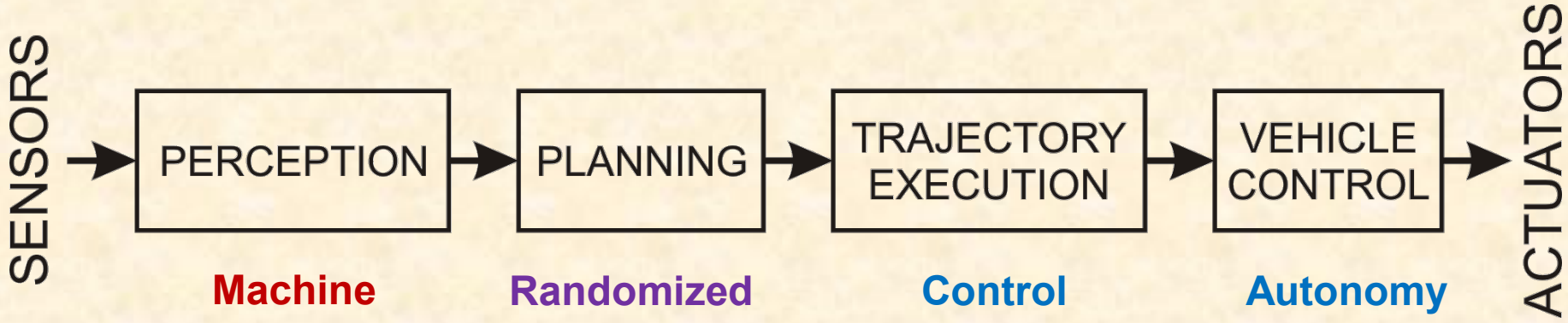


- Even though Newtonian Physics is useful
  - It requires accurate world model information (from perception??)

# Validating an Autonomous Vehicle Pipeline

Carnegie Mellon University




**Machine Learning Based Approaches**

➔ **Simulation & SOTIF approaches**

**Randomized & Heuristic Algorithms**

➔ **Run-Time Safety Envelopes**

➔ **Doer/Checker Architecture**

**Control Systems**

➔ **Control Software Validation**

➔ **Doer/Checker Architecture**

**Autonomy Interface To Vehicle**

➔ **Traditional Software Validation**

**Prediction & perception are uniquely difficult to assure**

# Importance of Behavior Prediction



- Free space: available drivable area
  - Move to where the free space is going to be
  - Can require fine grain classification


https://www.azquotes.com/quote/117311

# From Fail Silent to Fail Operational

Carnegie Mellon University

- Driver Assistance approach
  - Driver controls vehicle
  - Computers help
  - Fail silent computers
- ADS approach
  - Computer controls vehicle
  - Driver is out of the loop during operation
  - Computers keep working after a failure ("fail operational")
    - At least long enough for driver to take over in Level 3
    - More redundancy than conventional vehicle
    - Different fault management (e.g., pull to side of road)

UA 328 Feb 2021 https://bit.ly/3dPQRXZ

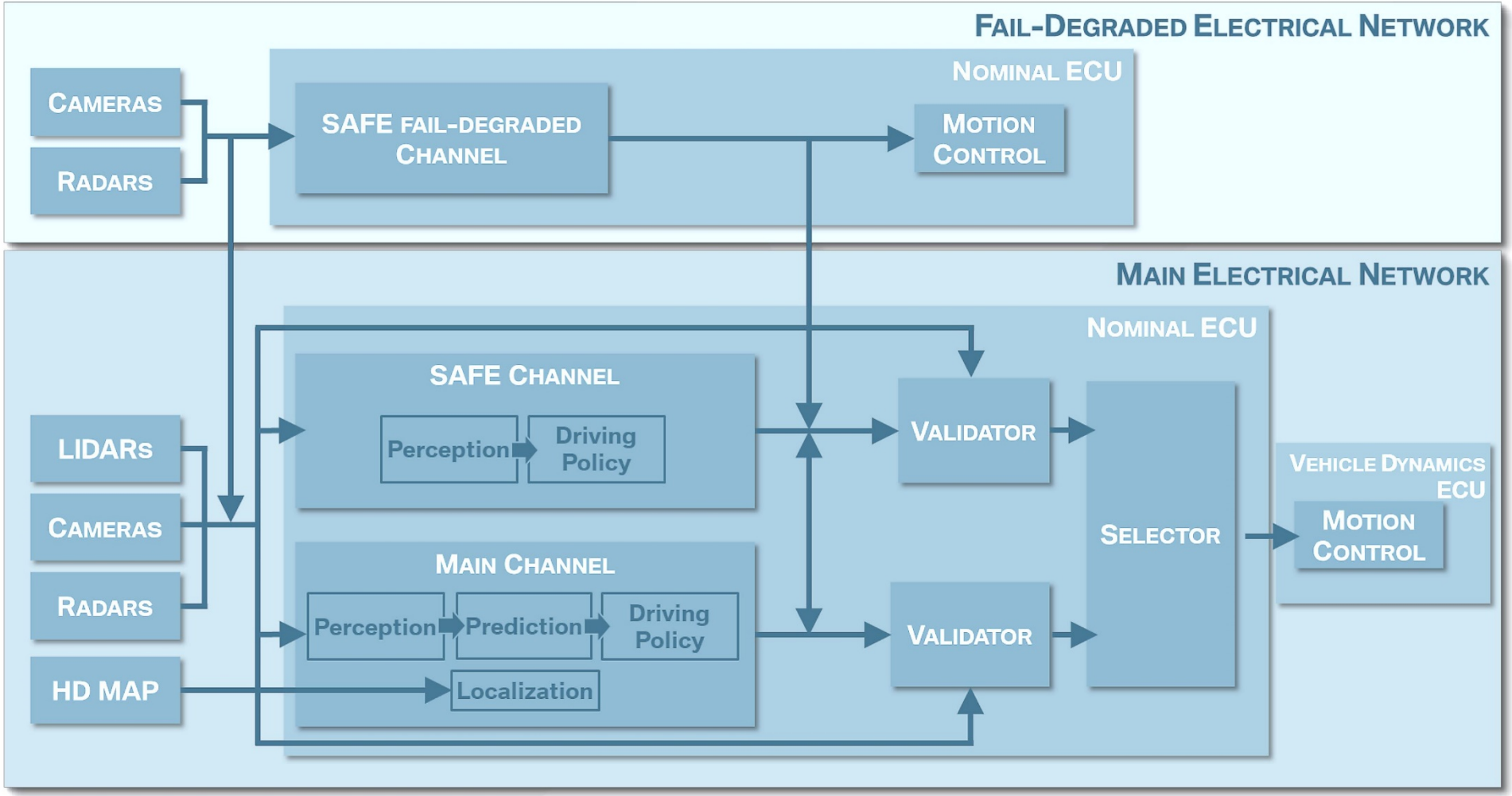



Figure 24. Implemented Redundancy Concept in the BMW ADS.

BMW VSSA https://bit.ly/3gCiiGw

# Redundancy & Decomposition

Carnegie Mellon University

- ASIL B(D) redundancy strategy:
  - Two ASIL B channels for net ASIL D
  - Failure independence required!


- Mitigate potential common cause failures:
  - Same perception/sensor fusion/planning algorithms
  - Same operating system, compiler, libraries, …
  - Same CPU types, network chips, discrete components, …
  - Same hardware boards (thermal; EMC; power distribution)
- Attaining high diversity (>90%) is difficult!
  - Requires significant, dedicated engineering effort

# Move To Centralized Architecture

Carnegie Mellon University

- Older architecture
  - ECU per major function
  - 1st Tier supplier does HW + SW + integration for ECU
- Newer architecture
  - Central computing ECU
    - Sensor fusion + path planning + vehicle control
    - Other functionality as well
  - Supplier + OEM software on same ECU
- Multi-function and multi-vendor software integration
  - Resource & functionality conflict management by OEM


https://bit.ly/3vo4zr7

# Changing Computing Architecture

- Feature specific ECUs ➔ centralization
- Fail silent ➔ fail operational strategy
- Significant effort on redundancy+diversity

Carnegie Mellon University