Scalable Methods for Verifying Autonomous Cyber-Physical Systems

Dr. Rahul Razdan

Rahul@razinstitute.com

John von Neumann Distinguished Fulbright

Széchenyi István University, Gyor, Hungary

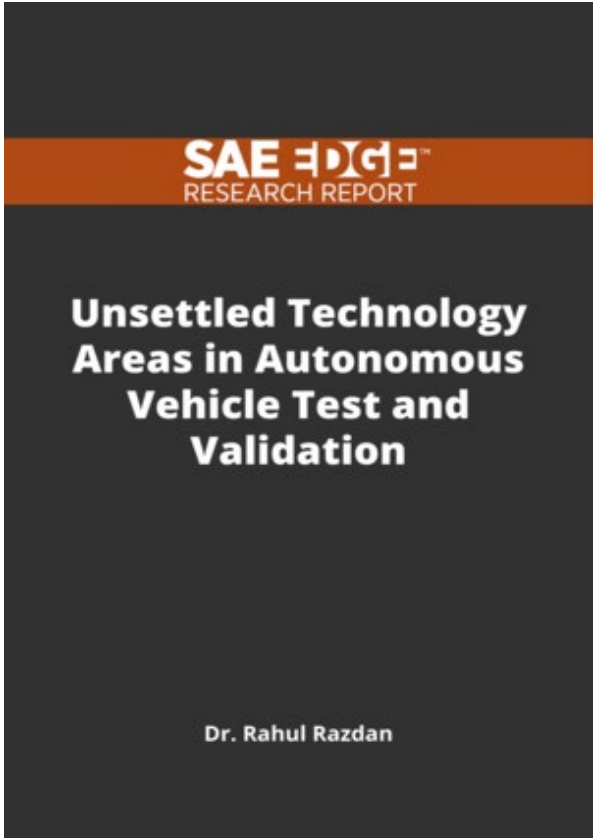
Background

Rahul: CPU Design ⇒ EDA ⇒ Startups ⇒ Research Institute

Transportation: Florida Polytechnic University [2016]

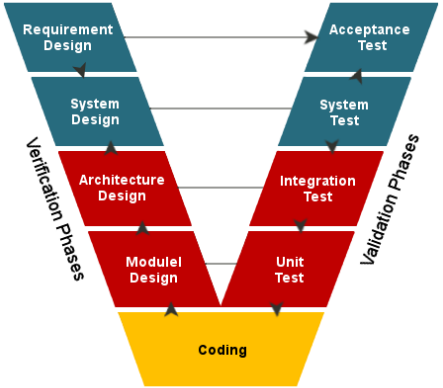
VENUE	FOCUS
Academic (IEEE, SAE)	AV Open-Source Environment: www.avvc.net
SAE Edge	Technology+Business+Government: book/course/research reports
Semi [SemiWiki, EPSnews]	Long-LifeCycle System Products (www.anew-da.ai)
Forbes Transportation	General Interest: education/humor/opinion

SAE Collaboration – Edge Reports

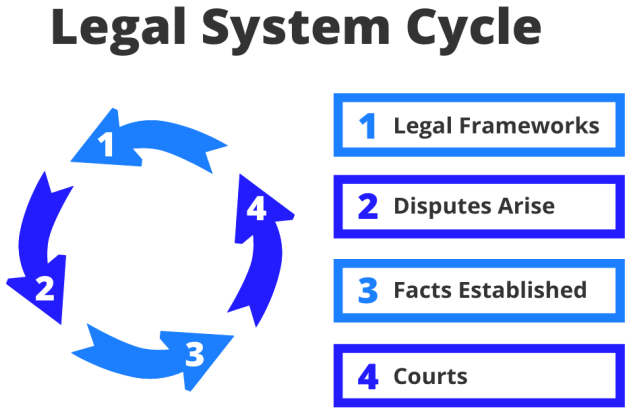


Contributor	Role
Dr. Avinash Balachandran	*Toyota Research Institute*
Satya Sreenivas	*IBM*
Dr. Raivo Sell	*Tallinn University of Technology, Estonia*
Dr. Dirk Langer	*Continental*
Jeff Lumina	*Jabil Circuit*
Nicholas Keel	*National Instruments*
Dr. M. Ilhan Akbas	*Embry-Riddle Aeronautical University*
Dr. Joachim Taiber	*Int'l Transportation Innovation Center (BMW)*
David Zuby	*Insurance Institute for Highway Safety*

The Basics (Part One)

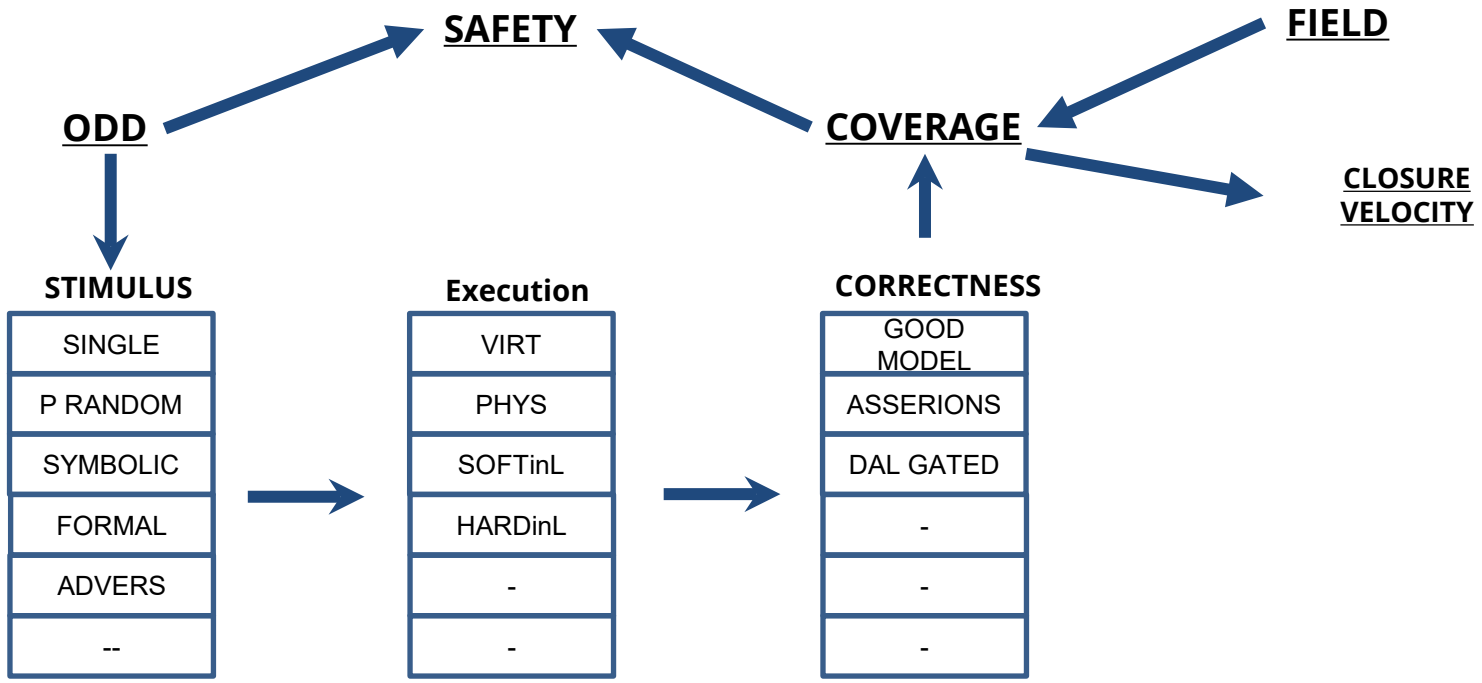


Verification/Validation

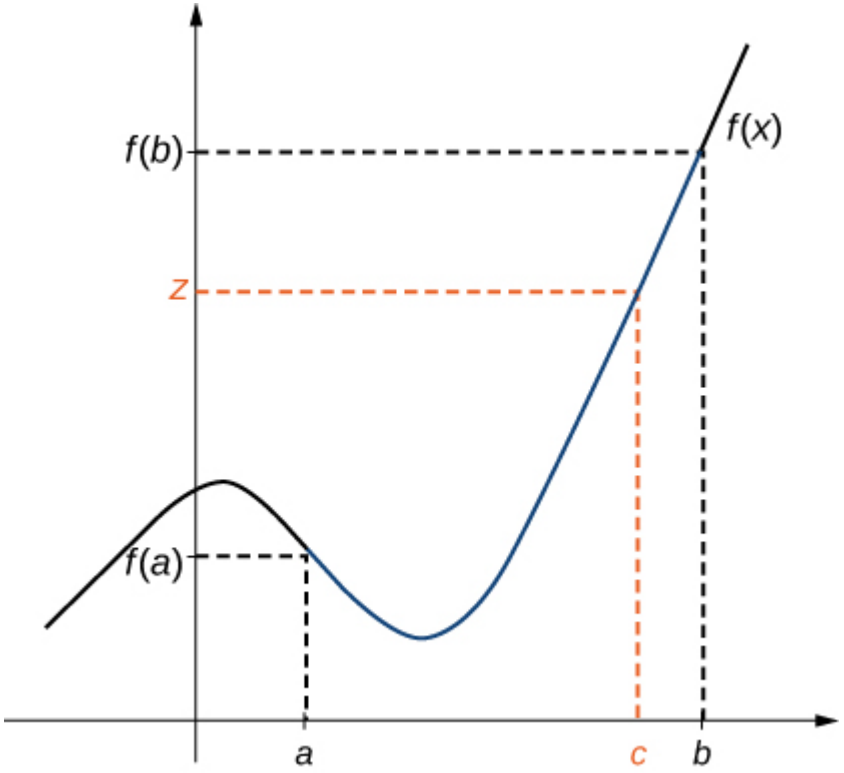


Validation/Safety/Liability

V&V FLOW


SAFETY
FIELD
ODD
COVERAGE
CLOSURE VELOCITY
STIMULUS
SINGLE
P RANDOM
SYMBOLIC
FORMAL
ADVERS
--
Execution
VIRT
PHYS
SOFTinL
HARDinL
-
-
CORRECTNESS
GOOD MODEL
ASSERIONS
DAL GATED
-
-
-

Problem Spaces (Physical Systems)

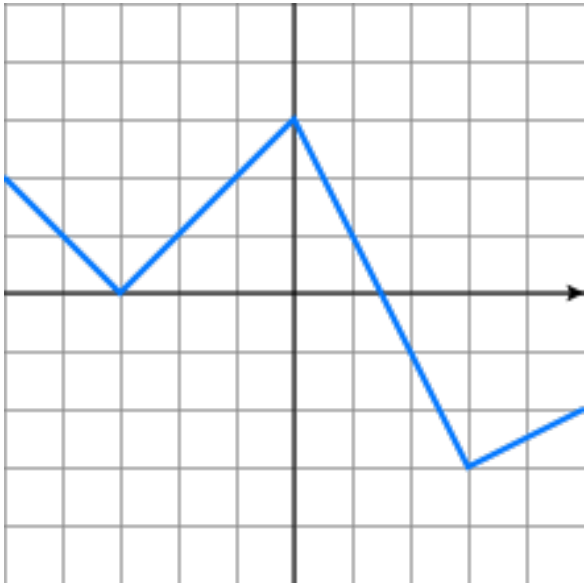


- Good Properties: Monotonicity and Continuous functions
- “Bad” Properties: Hyper Connected …worse case: butterfly effect

➔Rich history of validation using these properties in mechanical systems

➔Traditional Safety field developed with deep idea of risk assessment

Problem Spaces (Digital Systems)

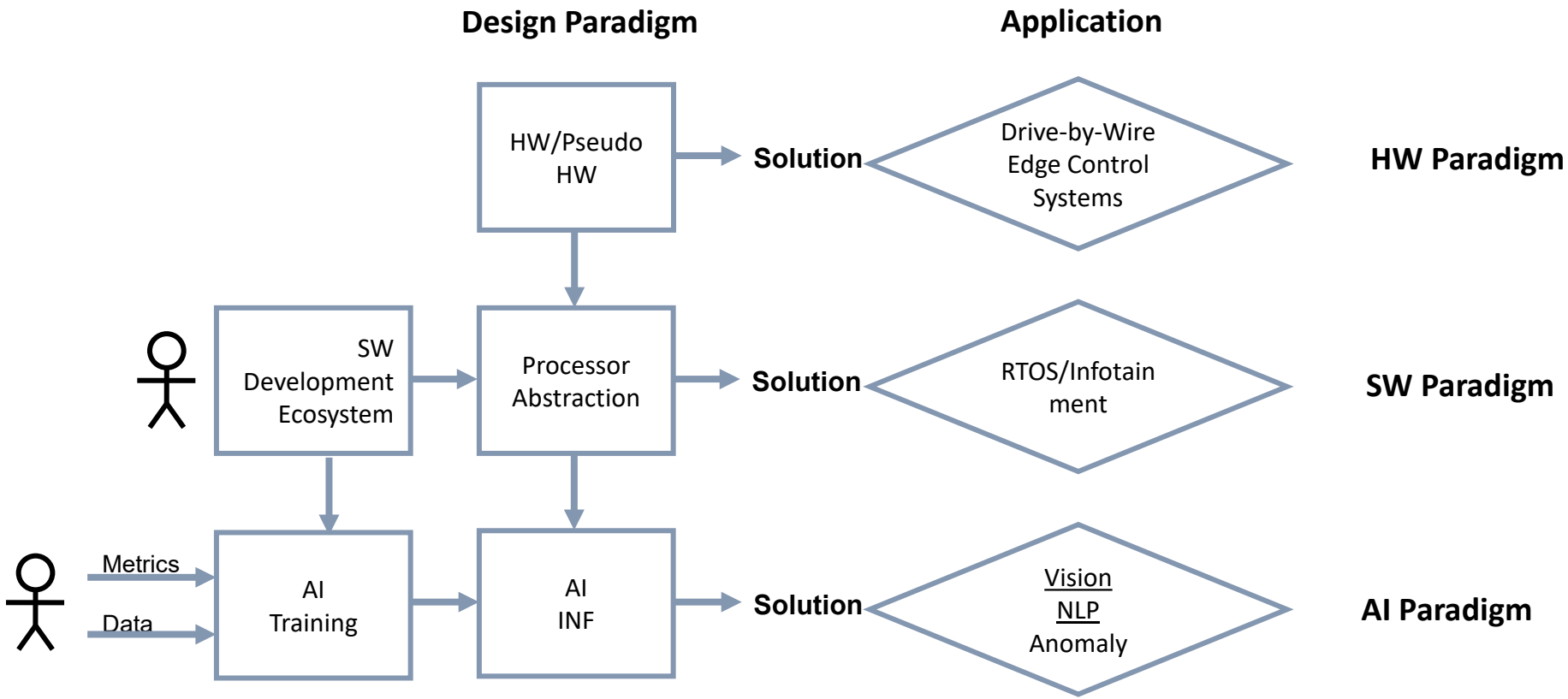


- Digital Systems (focus on Decision systems)
  - “Bad” Properties: not necessarily monotonic nor continuous
  - Good Properties:
    - Potential to unchain connections based on design+ODD+Thresholds
    - Potential to use abstraction to scale the process of validation
    - Controllability and Observability

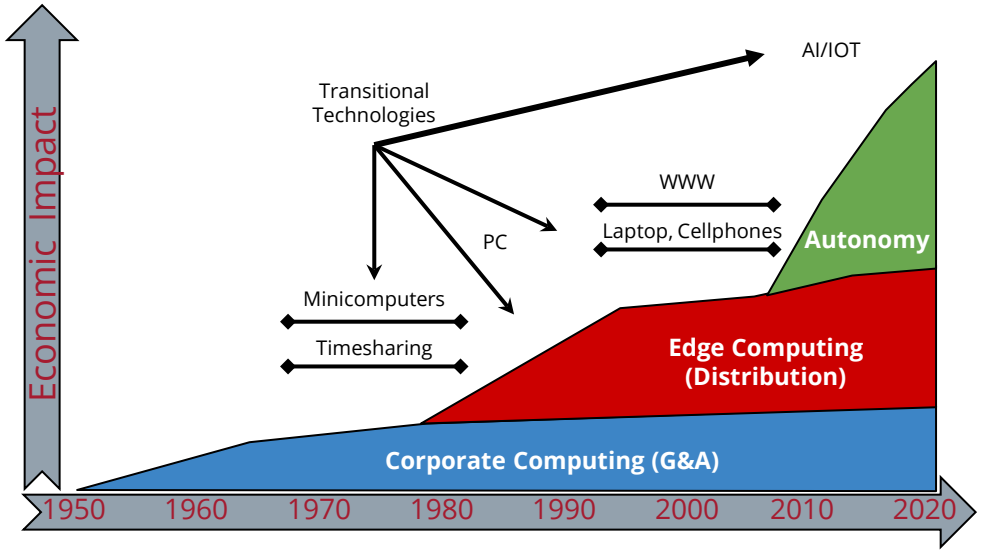
➔ Rich history of validation in current HW/SW flows

Autonomy: The devilish combination of Physical + Digital Decision Systems

AI/ML is the Next Big Abstraction



Why Is AI Important ?



Fundamental Technology Which May Enable the Next Set of Applications

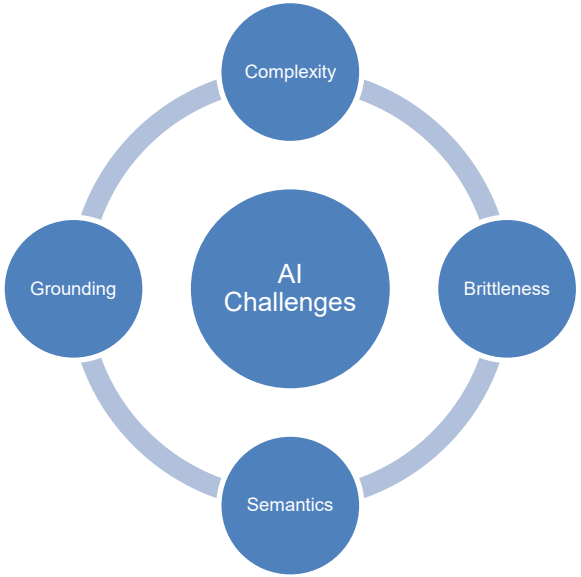
AI: The Quantum Physics of Computing

Conventional SW	ML Algorithms	Comment
Logical Theory	No Theory	ML algorithms can often just “work”.
Analyzable	Not Analyzable	SW Code vs ML Black Box
Causal	Correlation	The difference is important (optimization)
Deterministic	Non-Deterministic	ML algorithms are fundamentally probabilistic in nature.
Known Computational Complexity	Unknown Computational Complexity	For ML techniques, no generic method for computational complexity.

Forbes: Is Machine Learning The Quantum Physics Of Computer Science ?
Forbes: The Connection Between Astrology And Your Tesla AutoDrive

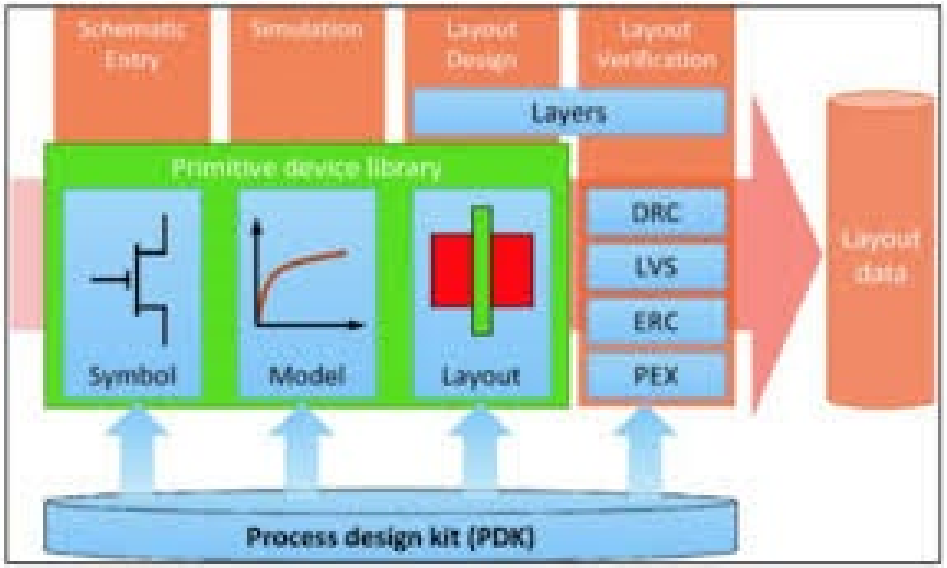
AI Challenges

- What is the right AI Model and why ?
  - Some theory.. Mostly empirical
- Will the model converge ?
  - No theory …
  - Addition, Vision, NLP, Astrology.. All look the same
  - Well behaved ➔ Brittle Models/Systems
- Robustness to noise ? How is this learned ? What is noise anyway ?

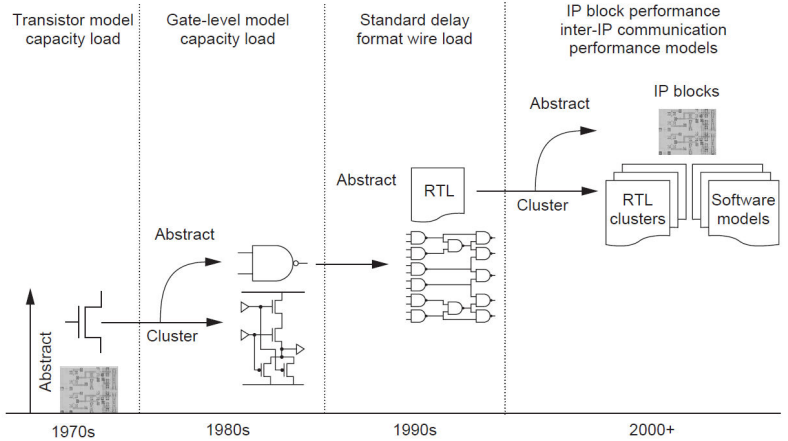


How do I know it works ?

Electronics/Semi V&V

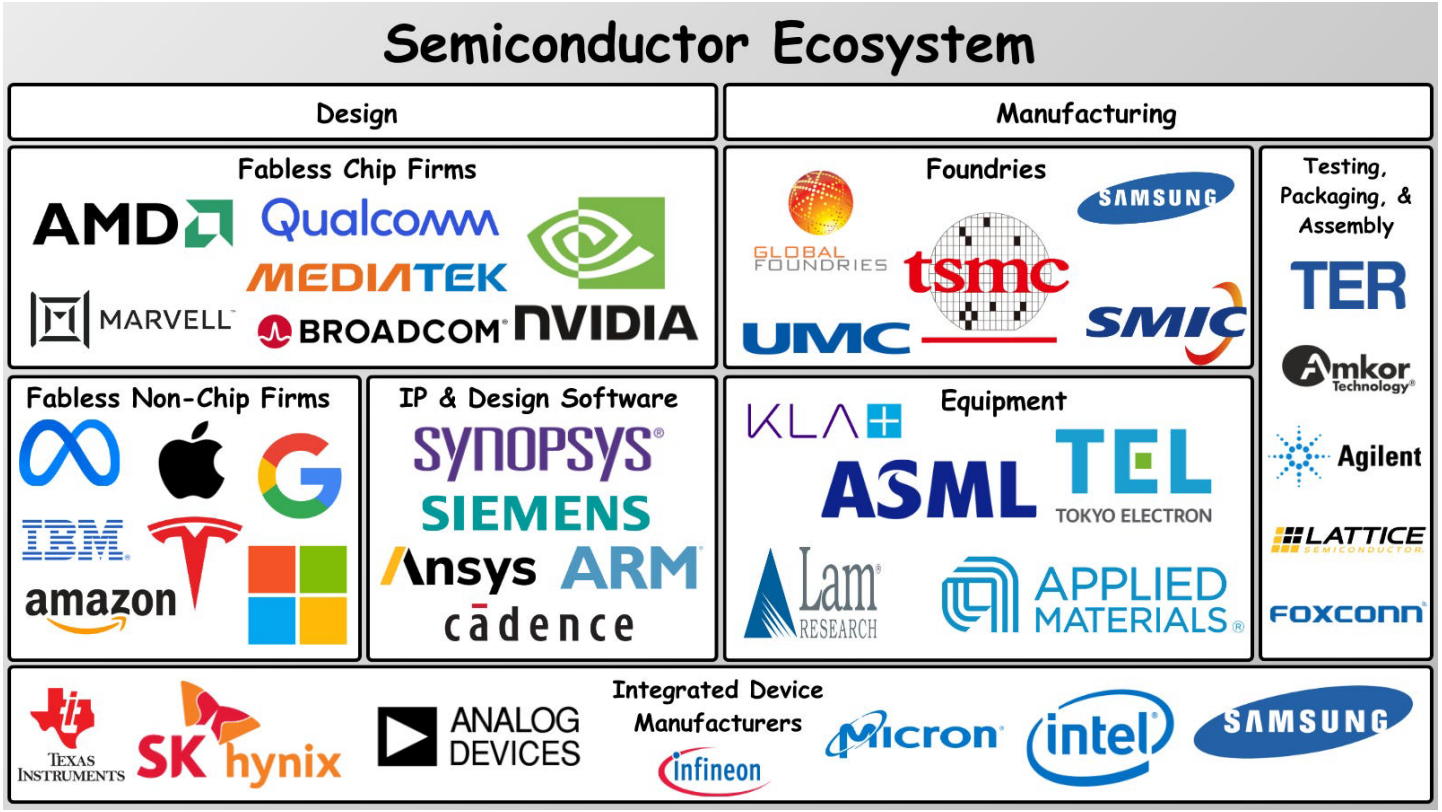


Physical to Virtual Mapping



Abstraction + Composition

Nested Design/Supply Chain ($2T+) Enabled by Abstraction

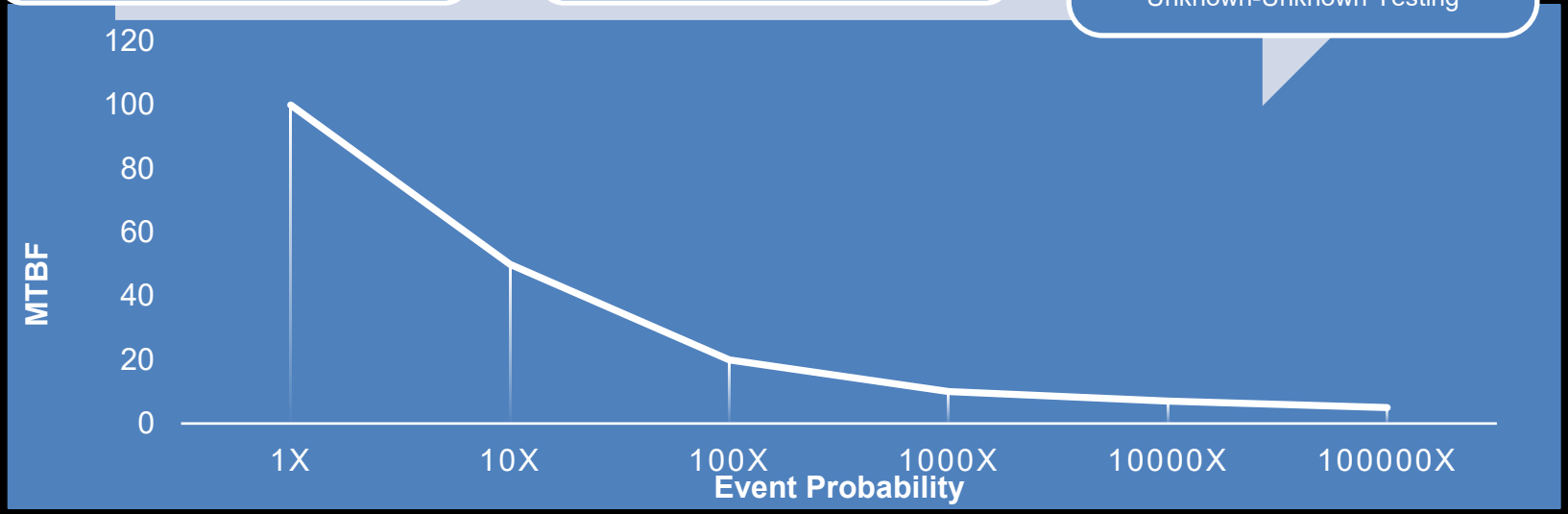


Scenario Test Gen

Research (Learning) Theory
Simulation
Known-Known Testing

Real World (Learning)
Test Track (Controlled Real World)
Known-unknown Testing

Real World (Production)Test Track (Diagnosis)
Test Case Acceleration (must be faster than real world)
Unknown-Unknown Testing



V&V Approaches/Challanges

- System Design Driven Flow:
  - AI components are SW+
  - AI vs SW (no structure, training data is the “program”)
  - Training Data Validation ? (noise vs data, ODD vs Training)
  - How to measure completeness ?
- Native AI Applications:
  - No System Spec… how to determine correctness ?
  - How to measure completeness ?
- V&V Toolkit:
  - System Spec when available or Anti-Specification if not (assertions)
  - Abstraction of Models and Tests
  - Coverage Buckets fractured by the Search Space
  - Adversarial AI Systems (ex collision avoidance)

There is a need for a research platform to accelerate the learning curve

AVVC – Open-Source Autonomy V&V Research Framework

AVVC Active Consortium Members

FLORIDA POLYTECHNIC UNIVERSITY

EMBRY-RIDDLE Aeronautical University

AVVC AUTONOMOUS VEHICLE VERIFICATION CONSORTIUM

TAL TECH

Berkeley UNIVERSITY OF CALIFORNIA

Extendibility

- Course Materials
- Algorithmic Enhancements
- Behavior / Prediction Analysis
- Sensors Compatibility
- Interference Issues

AVVC – Open-Source Autonomy V&V Research Framework

Critical Components:

- Captures a physical AV environment
- Performs extensive VV&C in a virtual environment
- Generate validation results and diagnostic data
- Provides a path back to physical testing

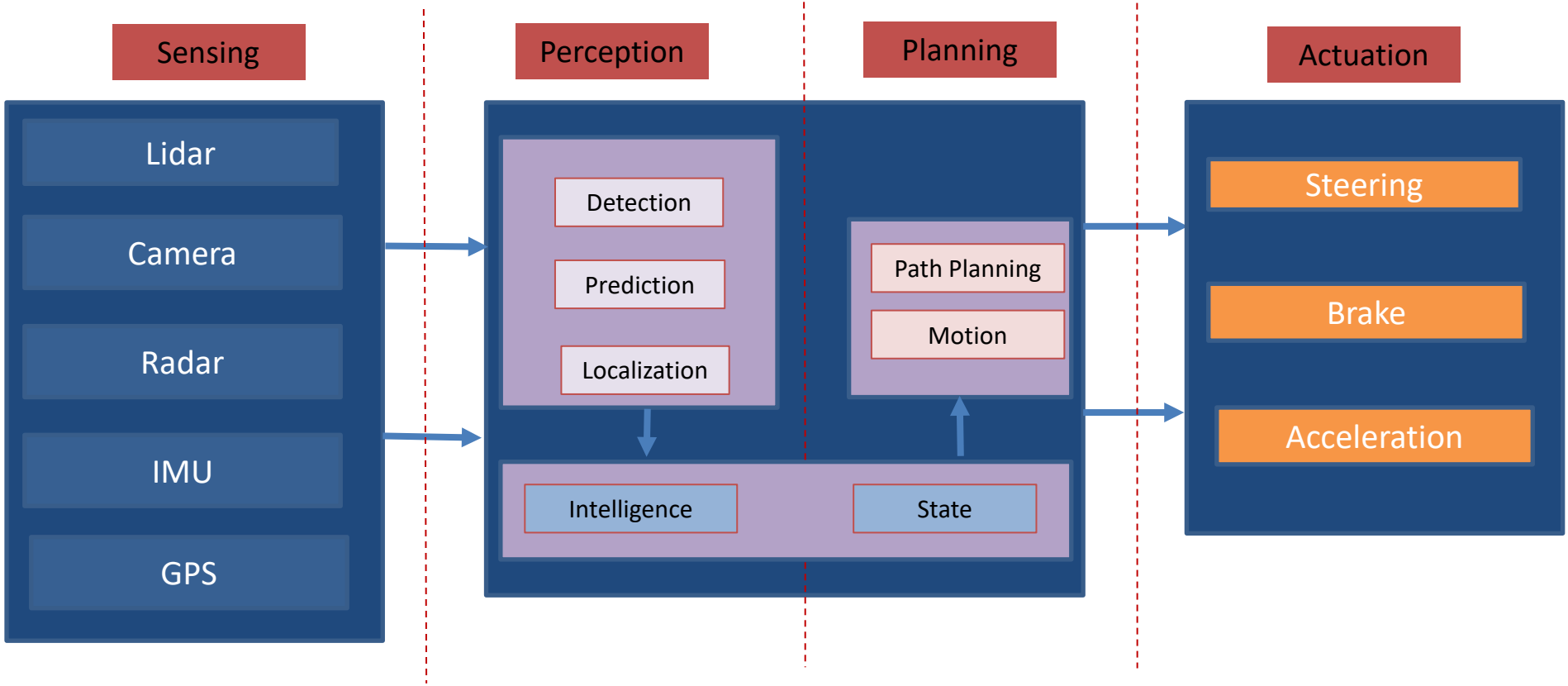
Leverage Existing Open-Source Components:

- Autosim
- SCENIC
- SUMO
- Autoware

AVVC Framework:

- ➢ Build coherent and integrated Design of Experiment (DoE) capability.
- ➢ Fixed-route autonomous public transportation as initial Use-Model

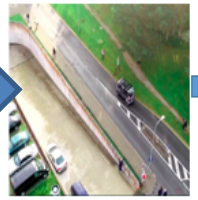
Autonomy Software Stack



AVVC – AVVC to Physical Track

Physical Test Track e.g. GoMentum Suntrax Test Track

Drone to capture the Point cloud data of a route



Point Cloud Data

Agisoft Metashape Tool

Process the point cloud data using Metashape to generate a Unity based 3D Scene

3D Scene Asset in Unity

Use the 3D Scene generated and annotate additional information with lanelet2 to generate exact 3D Map environment

Digital Twin Simulation

PolyVerif validation with different scenario's

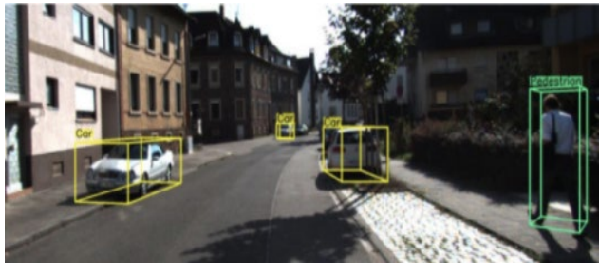
- AV Perception stack
- AV Control algorithm stack
- AV Path planning decisions
- Response times of the AV
- Diagnostic data for analysis
- High Level Summary reports

NCAP Scenario Testing: AEB + FCW Tests

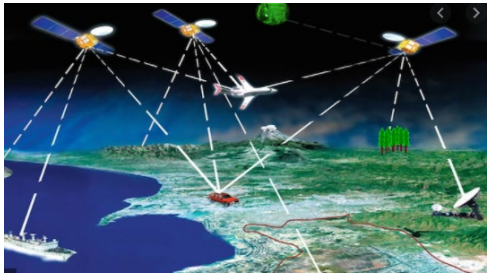
EuroNCAP-2018 AEB VRU	Scenario	Day/Night	Speed (60 kmh max)
CPFA-50	Car-to-Pedestrian Far side Adult 50%	Day	20-60 kmh
CPNA-25	Car-to-Pedestrian Nearside Adult 25%	Day	20-60 kmh
CPNA-75	Car-to-Pedestrian Nearside Adult 75%	Day	20-60 kmh
CPLA-25	Car-to-Pedestrian Longitudinal Adult 25%	Day	20-60 kmh
CPLA-50	Car-to-Pedestrian Longitudinal Adult 50%	Day	20-60 kmh
CBNA-50	Car-to-Bicyclist Nearside Adult 50%	Day	25-60 kmh
EuroNCAP-2020 AEB VRU			
CBNAO-50	Car-to-Bicyclist Nearside Adult Obstructed 50%	Day	25-60 kmh
CBFA	Car-to-Bicyclist Farside Adult 50%	Day	25-60 kmh
CPRA	Car-to-Pedestrian Reverse Adult 50%	Day	4,8 kmh
CPTA-50	Car-to-Pedestrian Turning Adult 50%	Day	10,15,20 kmh
EuroNCAP-2019 Lane Support Systems			
ELK Road Edge	Emergency Lane Keeping - Road Edge	Day	25-60 kmh

Test critical scenario identified from PolyVerif Suite on Physical Track

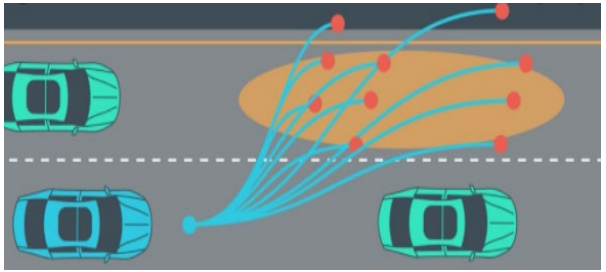
AVVC – Open-Source Autonomy V&V Research Framework



- **Detection Validation:** Do the sensors actually “see” the objects of interest ?
- **Perception Validation:** Having “detected” the objects, are they recognized sufficiently to determine future movement?

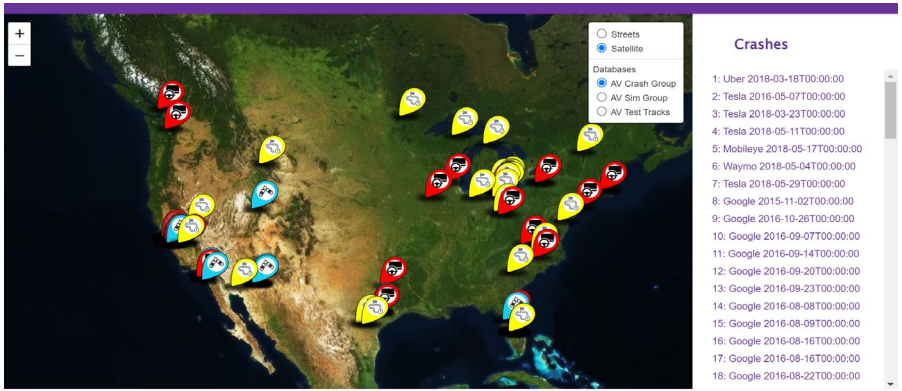


- **Location Validation:** Decisions on movement are based on current position, is the current position “known” ..both globally and relative to local objects.



- **Decision Validation:** Even when perception is perfect and control systems provide stability, are the correct choices on path planning being made?
- **Control Validation:** Many tasks in autonomy are control systems (e.g. Cruise control). Are these systems stable under environmental noise?

Test Scenario Formulation

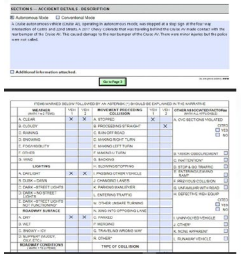


foretellix
verifying autonomous systems

TAL TECH

Berkeley
UNIVERSITY OF CALIFORNIA

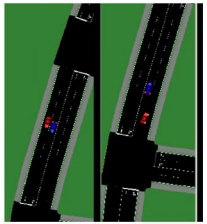
Functional



Logical

```
scenario dut.bl_1x2:
    npc: car
# ... junction: junction
    p_dur: time
    p_dut_speed_start: speed
    p_npc_speed_at_dut_start: speed
    dut_path: path_over_junction
    npc_path: path_over_junction
    keep(dut_path.junction == npc_path.junction)
    keep(dut_path.in_road != npc_path.in_road)
    npc_to_dut_time_delta: time with:
        keep(default it in [(-1)..1]second)
    dut_direction: direction with:
        keep(default it in [left,straight,right])
    npc_direction: direction with:
        keep(default it in [straight])
    do cross_junction_s: serial(duration: p_dur):
        mix():
            dut_traverse: dut.car.traverse_junction(dut_path)
            npc_traverse: npc.traverse_junction(npc_path)
        with:
            synchronize(slave: npc_traverse.enter.end,
                        master: dut_traverse.enter.end,
                        offset: npc_to_dut_time_delta)
```

Concrete



- Developed scenario database using existing AV accidents
- Worked with industry on scenario generation and analysis
- Analyzed NHTSA accident database
- Working with UC Berkeley on formal scenario description and integration
- Working with TalTech on integration of simulation and Naturalistic Field Operational Tests

EMBRY-RIDDLE
Aeronautical University

Case Study: JTA Project

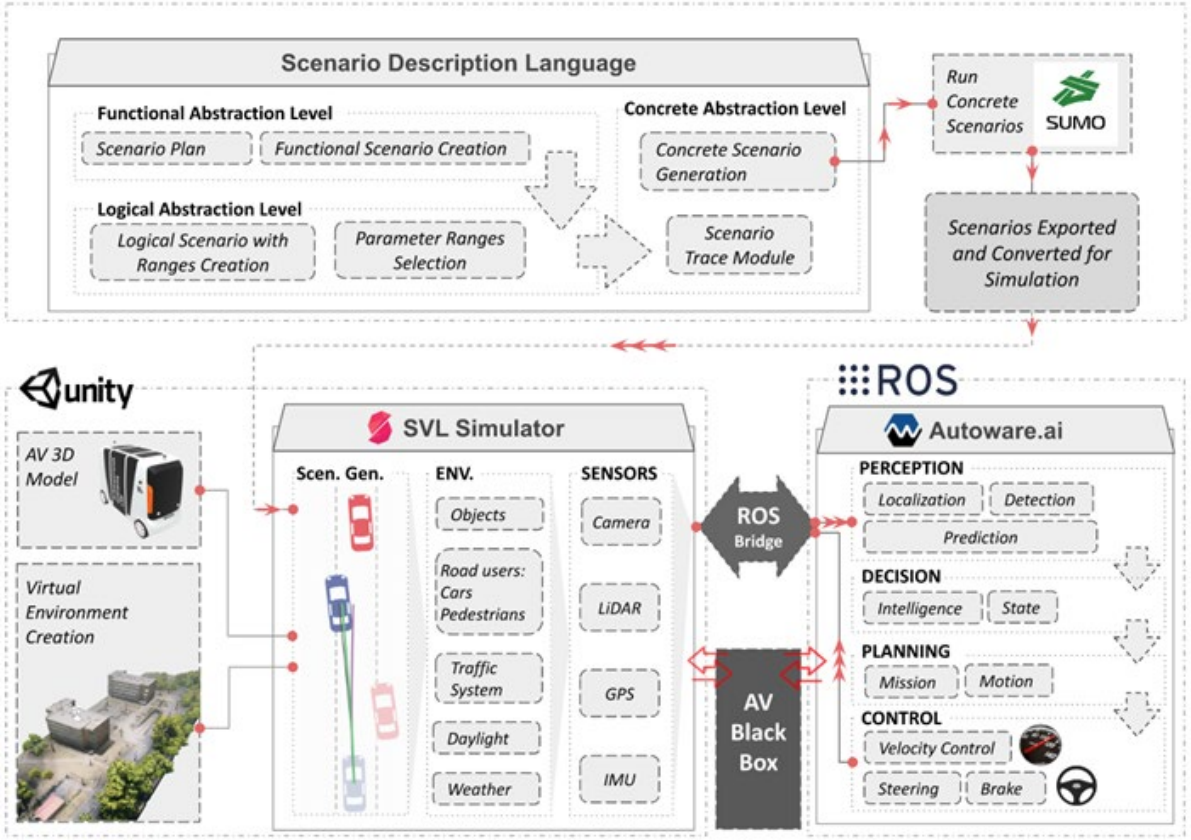
- Creation of digital twin for simulation of JTA routes

- Current digital twin have following features
  - Road curvature, Junctions
  - Similar building, Side objects architectures

- Creation of interesting test cases
  - Weather and traffic simulation
  - Round about scenarios
  - Lane merging
  - Blind spots
  - Pedestrian collision
  - Road incidents

- The JTA Digital Twin and Test Scenarios created could be utilized for validation of different AV stacks and hardware setups.

Case Study: TalTech



Summary (Driving Research Questions)

- Cyber-Physical Systems:
  - What are the right abstractions which leads to “separation of concerns ?”
  - What are design simplifications which can lead to an interesting theory of composability ?
  - How does one define completeness ?
- AI Component Special Challenges: (causality, determinism, etc)
- What is the next level of Safety theory and how does it connect to legal liability ?

RAZDAN
RESEARCH INSTITUTE

The Basics Part 2

- Operational Design Domain (ODD) ➔ environmental constraints
  - Argument for Completeness (coverage)
  - Closure Velocity on defects

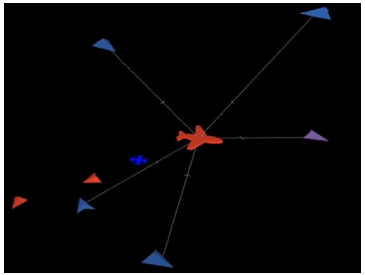
- Divide and Conquer Techniques
  - Functional Decomposition (subcomponets, integration cycle)
  - Abstraction Construction (verify abstractions separately, stich together and patch)

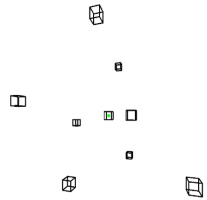
Aerial Domain

- Scenario generation methodology is modular and implementation agnostic

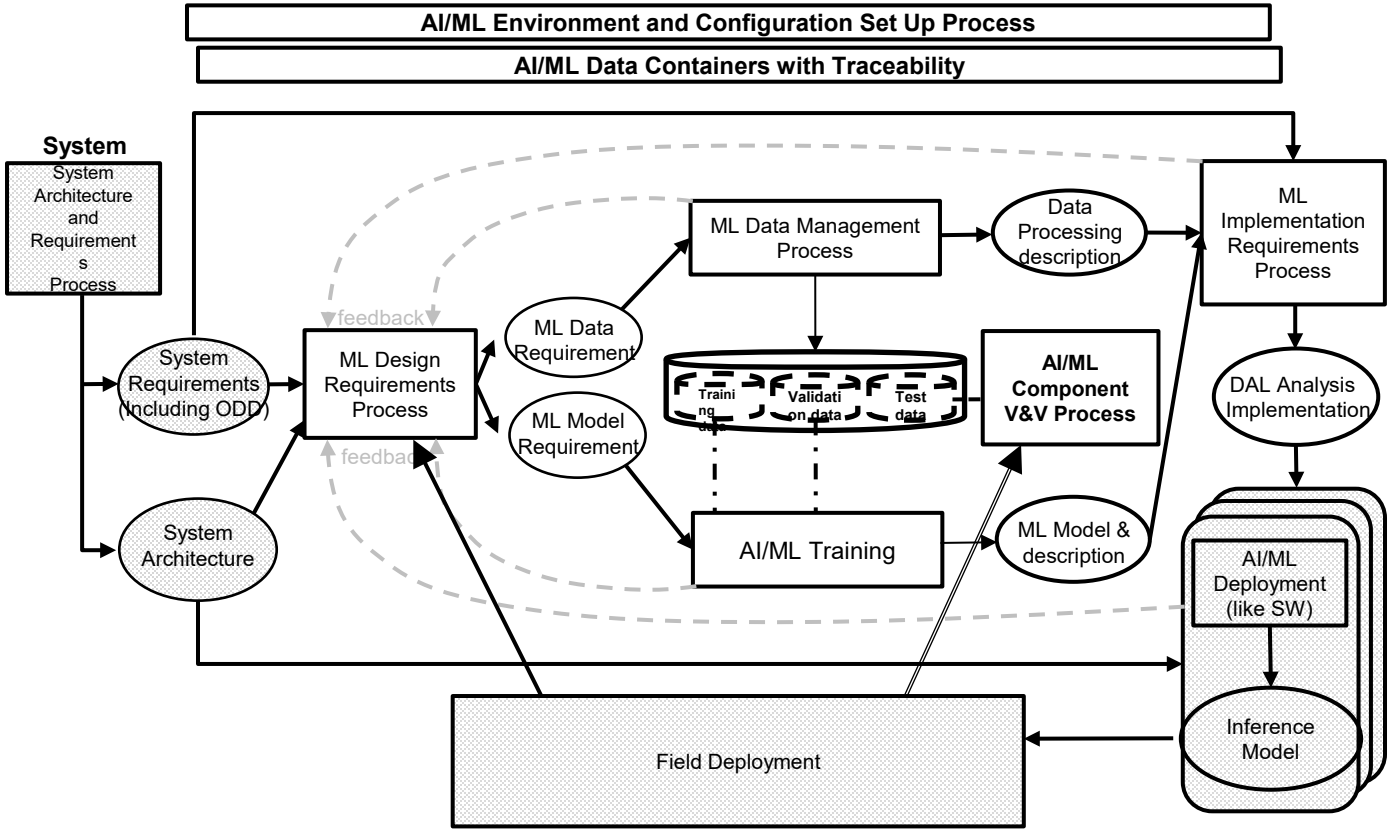
- Adaptation for UAS validation

- Adopting multiple simulators
  - Multi-agent simulation for abstract scenarios
  - Game-engine based simulation for visual scenarios
  - Network simulation for communication-based scenarios
  - ArduPilot and PX4 integration



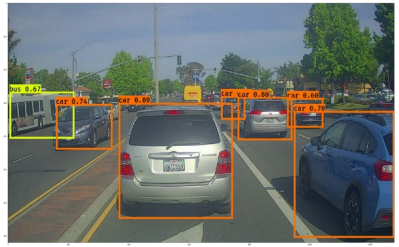


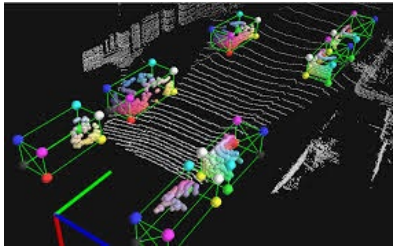
AI inside of System Design ? (FAA G34)



AVVC – Validation Layers

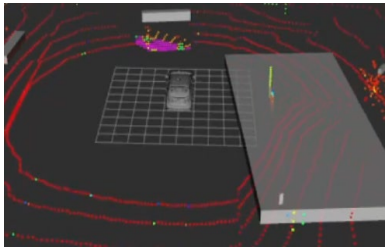
Perception Validation





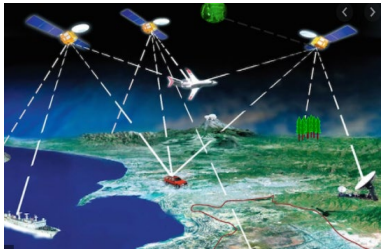
Reports :
- Object Detection Range in Simulator and Autoware
- Object Detection : Success
- Range Detection Rate

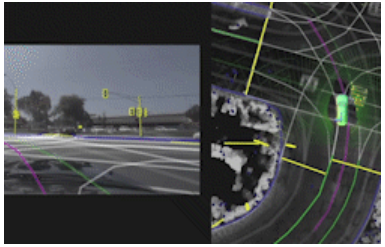
Control Validation



Reports:
- Time-to-Collision Calculations
- Response Time in Simulator & Autoware
- Delay in response time
- Control

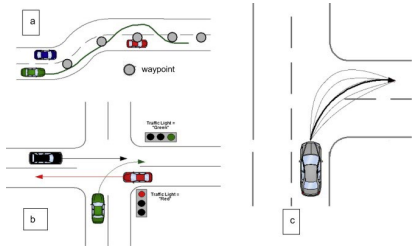
Localization Validation





Reports:
- Per frame deviation
- Max/Min/Mean Deviation

Mission Planning Validation

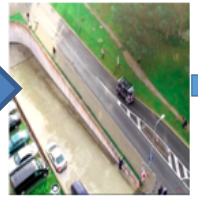


Reports:
- Mission Completion Statistics
- Obstacle Avoidance Testing Report
- Controls Testing Reports

Abstraction: Critical Questions

Physical Test Track e.g. GoMentum Suntrax Test Track

Drone to capture the Point cloud data of a route



Point Cloud Data

Agisoft Metashape Tool

Process the point cloud data using Metashape to generate a Unity based 3D Scene

3D Scene Asset in Unity

Use the 3D Scene generated and annotate additional information with lanelet2 to generate exact 3D Map environment

Digital Twin Simulation

PolyVerif validation with different scenario's

- AV Perception stack
- AV Control algorithm stack
- AV Path planning decisions
- Response times of the AV
- Diagnostic data for analysis
- High Level Summary reports

NCAP Scenario Testing: AEB + FCW Tests

EuroNCAP-2018 AEB VRU	Scenario	Day/Night	Speed (60 kmh max)
CPFA-50	Car-to-Pedestrian Far side Adult 50%	Day	20-60 kmh
CPNA-25	Car-to-Pedestrian Nearside Adult 25%	Day	20-60 kmh
CPNA-75	Car-to-Pedestrian Nearside Adult 75%	Day	20-60 kmh
CPLA-25	Car-to-Pedestrian Longitudinal Adult 25%	Day	20-60 kmh
CPLA-50	Car-to-Pedestrian Longitudinal Adult 50%	Day	20-60 kmh
CBNA-50	Car-to-Bicyclist Nearside Adult 50%	Day	25-60 kmh
EuroNCAP-2020 AEB VRU			
CBNAO-50	Car-to-Bicyclist Nearside Adult Obstructed 50%	Day	25-60 kmh
CBFA	Car-to-Bicyclist Farside Adult 50%	Day	25-60 kmh
CPRA	Car-to-Pedestrian Reverse Adult 50%	Day	4,8 kmh
CPTA-50	Car-to-Pedestrian Turning Adult 50%	Day	10,15,20 kmh
EuroNCAP-2019 Lane Support Systems			
ELK Road Edge	Emergency Lane Keeping - Road Edge	Day	25-60 kmh

Test critical scenario identified from PolyVerif Suite on Physical Track