MITIGATING
DESIGN ERROR ARCHETYPES
IN THE DEVELOPMENT OF EXPLAINABLE-MACHINE LEARNING (X-ML) SYSTEMS

Autonomous Shuttle Bus Accident

Lance Sherry, Jim Baldo, Brett Berlin, Oleksandra Snisarevska-Donnelly

AI-4-SE

Oct 28 03:30 – 4:00 pm

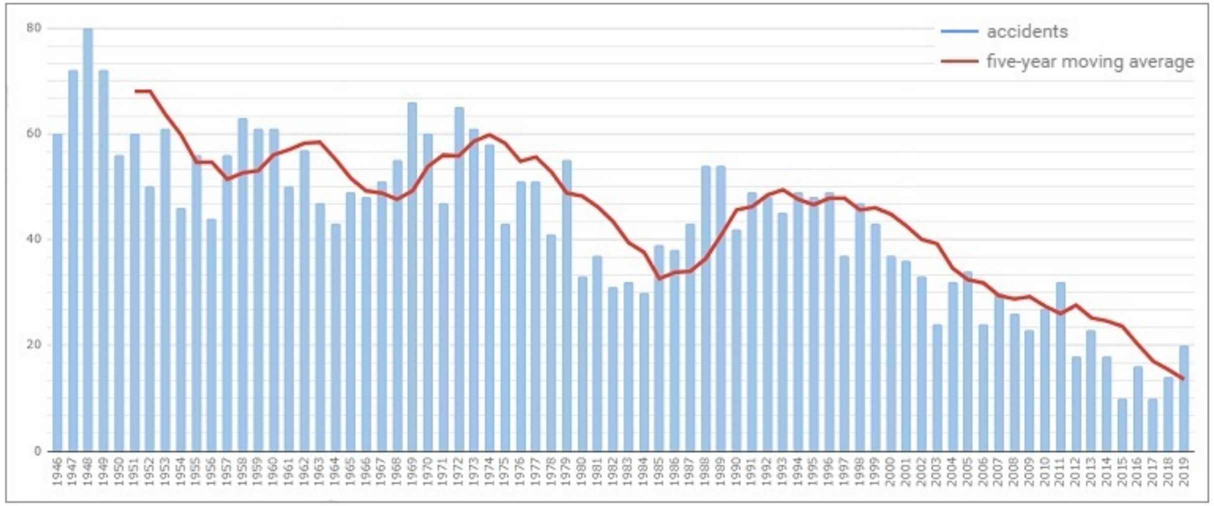
lsherry@gmu.edu

Center for Air Transportation Systems Research at George Mason University

Table of Contents

Motivation

Airline Accidents (1946 – 2019)



Modern flight deck (high levels of autonomy)

Sophisticated safety-culture/safety management system

Road to Zero: A Plan to Eliminate Roadway Deaths

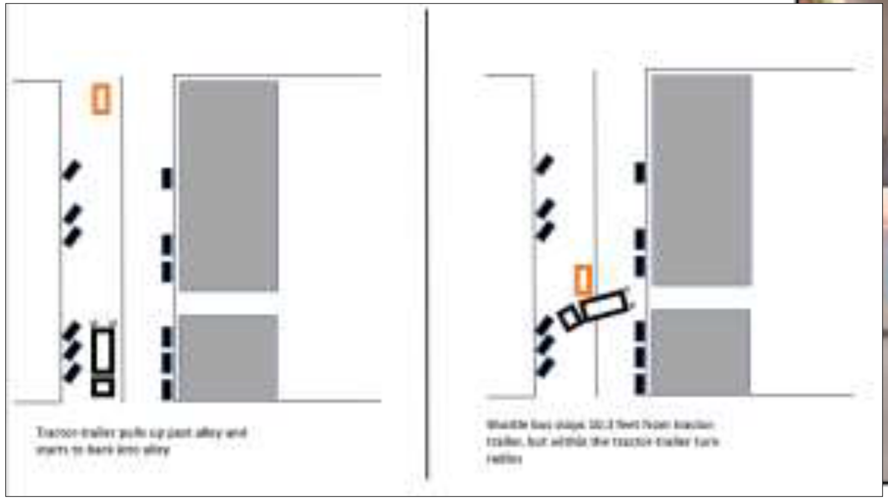
ROAD TO ZERO

Road to Zero Coalition Steering Group

Motivation

Nov 8, 2017 at 12:07pm

NTSB Report: Low-Speed Collision Between Truck-Tractor and Autonomous Shuttle, Las Vegas, Nevada, November 8, 2017
https://www.ntsb.gov/investigations/AccidentReports/Reports/HAB1906.pdf



Motivation

NTSB Probable Cause:

- **"the truck driver's action of backing into an alley, *and his expectation that the shuttle would stop at a sufficient distance from his vehicle to allow him to complete his backup maneuver*"**

- Design did not include corner-case
  - Tractor-trailer backing up with turn radius
- Test cases also missing this situation

Motivation:

NTSB Contributing Factor

“attendant not being in a position to take manual control of the vehicle in an emergency”

- Attendant role an “afterthought”
- Not explicit design of procedures or user-interface
- Aviation requires definition of Emergency Procedures (and re-current training)

Table of Contents

Research Question

- What <u>Design Errors</u> can occur with X-ML design of Operationally Embedded Control System?

  - Accidents/Incidents caused by **<u>Op Embedded Control System</u>**
    - Inappropriate Actuator Commands from Operationally Embedded Control System
      - Equipment Malfunctions vs Design Errors

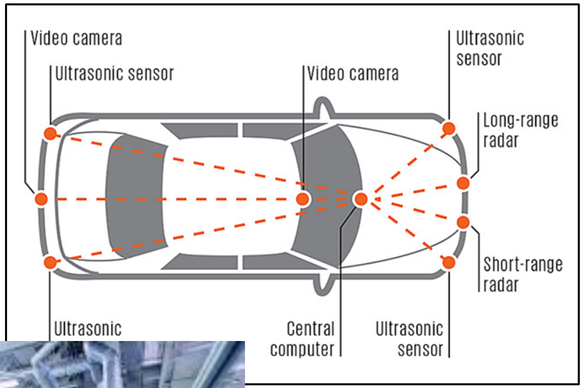
Table of Contents

Operationally Embedded Control Systems

- Embedded on vehicle or plant
- Provide Guidance and Control functions to perform Mission
- Complex
  - Over 100 input signals
  - Over 10 actuator command outputs

Operationally Embedded Control Systems

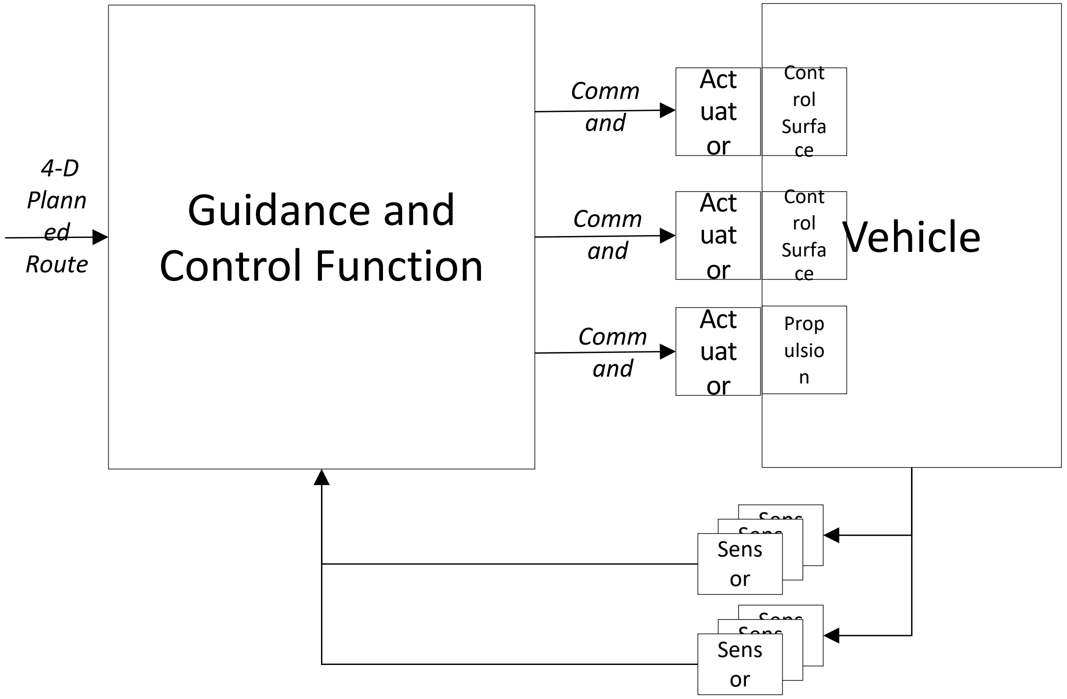
- Examples:
  - vehicle navigation systems
  - robotics
  - processing “plant” control
  - power generation, transmission, distribution management
  - expert decision support systems
    - Health care
    - Legal advice
    - Finance
    - Trading
    - …



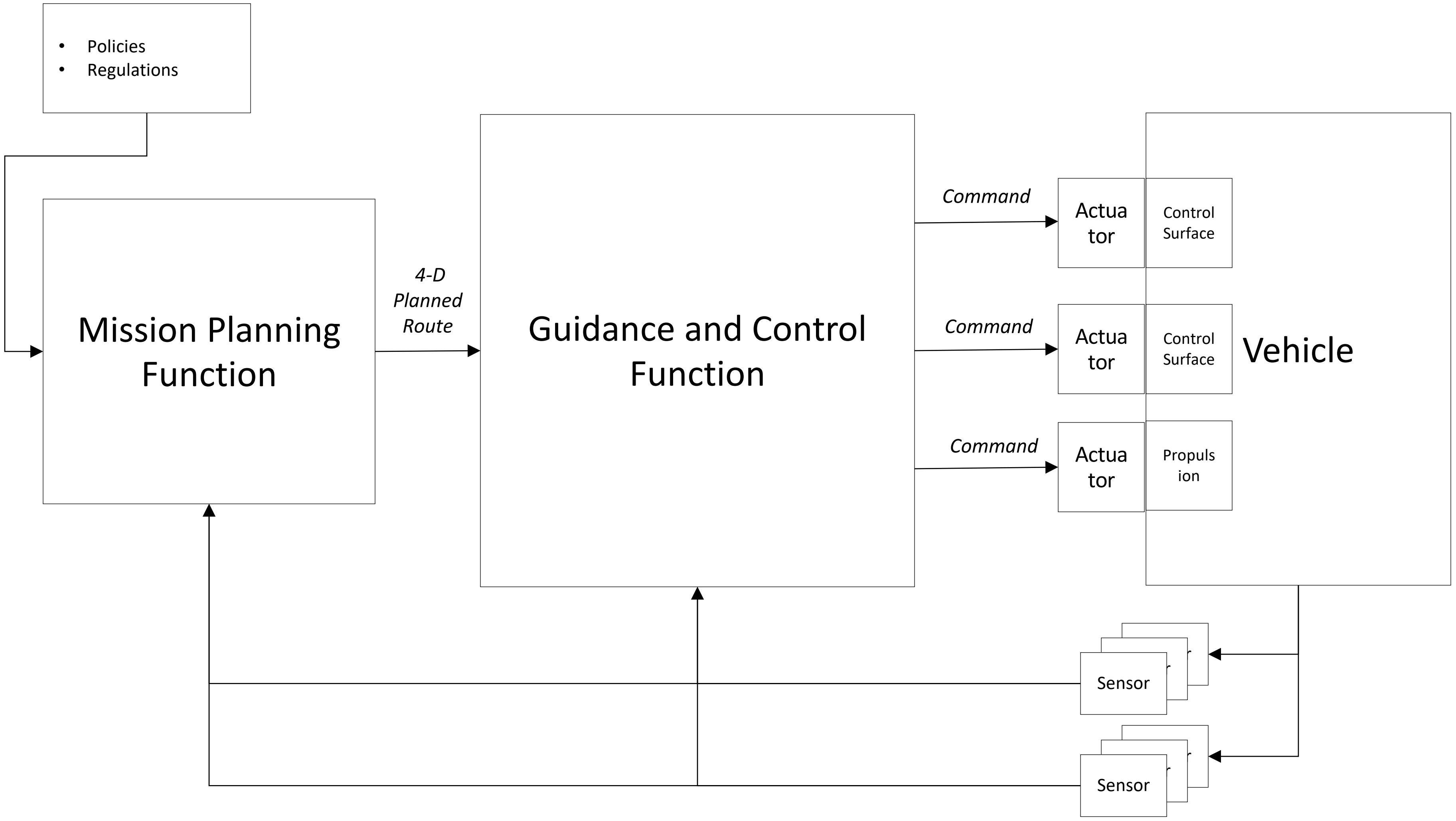
Operationally Embedded Control Systems

- perform complex real-time decision-making based on emerging situations in the environment

- Stimulus-Response
  - In real-time
  - Emerging situations in Mission
  - Meet Mission objectives
  - Manage normal & abnormal situations



Example: Vehicle Guidance and Control Function

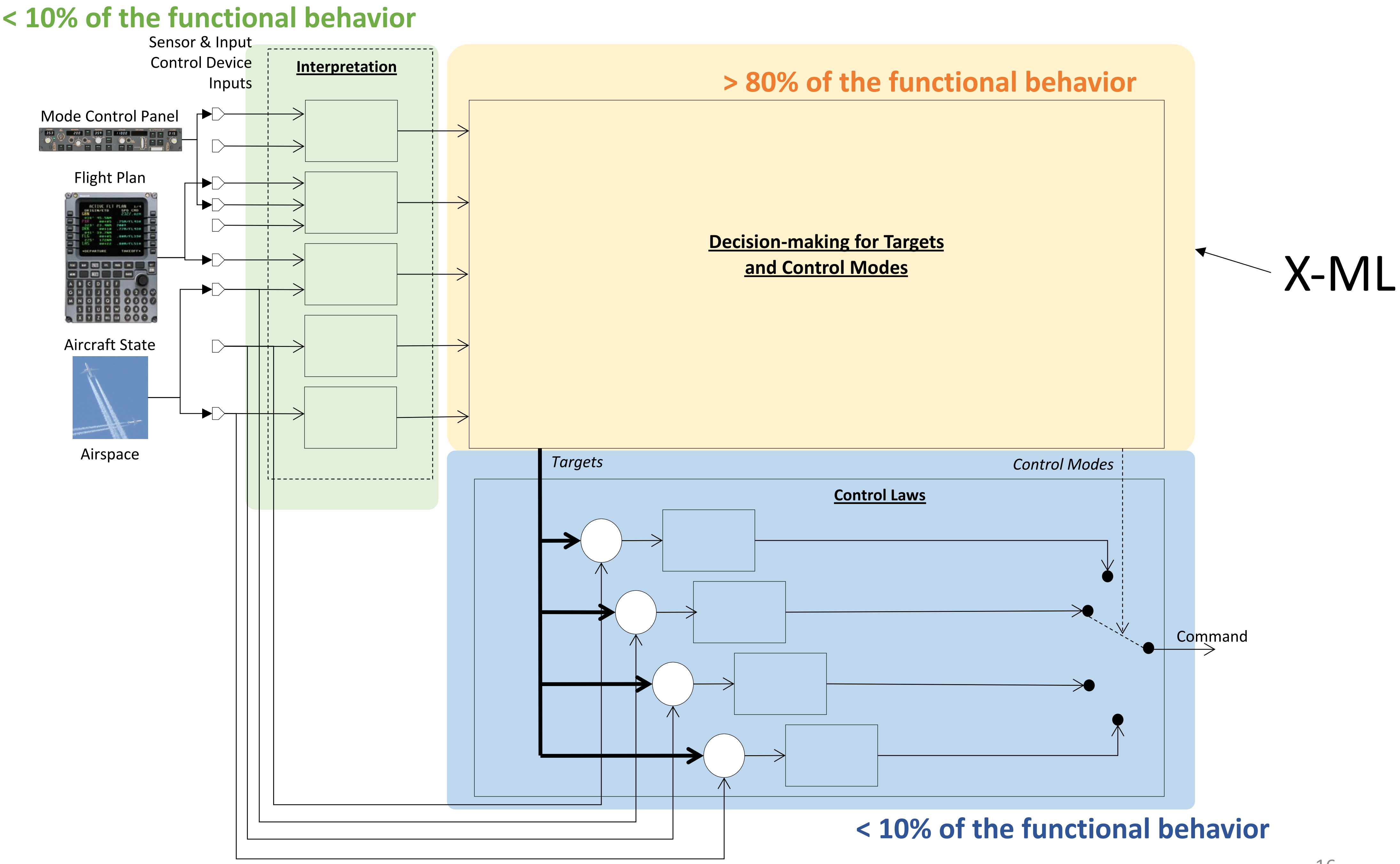


Example: Vehicle Guidance and Control Function

G&CF (Inputs, Outputs)	Fixed Wing	Automobile
4-D Planned Route	"Flight plan" • 4-D • Navigation Procedures • Air Traffic Control • Traffic avoidance • Terrain avoidance • Env. – Windshear	"Route" • 4-D • Roadway Rules • Signage and Traffic Lights • Traffic avoidance • Terrain avoidance • Env. – surface conditions, visibility
Commands	• Elevator • Aileron • Rudder • Thrust	• Accelerator/Brake • Steering

Example: Vehicle Guidance and Control Function

- Real-time Stimulus-Response
- Operational "smarts" to complete the Mission
- Three components:
  1. Control Laws
     - Closed-loop control laws (continuous mathematics)
     - Designed based on models of vehicle and actuator dynamics
  2. Decision-making for Targets and Control Modes
     - Decision (logic)
     - Designed based on:
       - Closed-loop control law operational boundaries
       - Vehicle performance operational limits
       - Mission operational rules and constraints
  3. Interpretation
     - Translate sensor/user-interface input data into operationally meaningful mission data



< 10% of the functional behavior
Sensor & Input Control Device Inputs
Interpretation
> 80% of the functional behavior
Mode Control Panel
Flight Plan
Aircraft State
Airspace
Decision-making for Targets and Control Modes
X-ML
Targets
Control Modes
Control Laws
Command
< 10% of the functional behavior

Operationally Embedded Control System

Definition of Terms:

- Interpretation
- Decision-making
- Control Laws

- Inputs/States
- Targets
- Control Modes
- Actuator Command

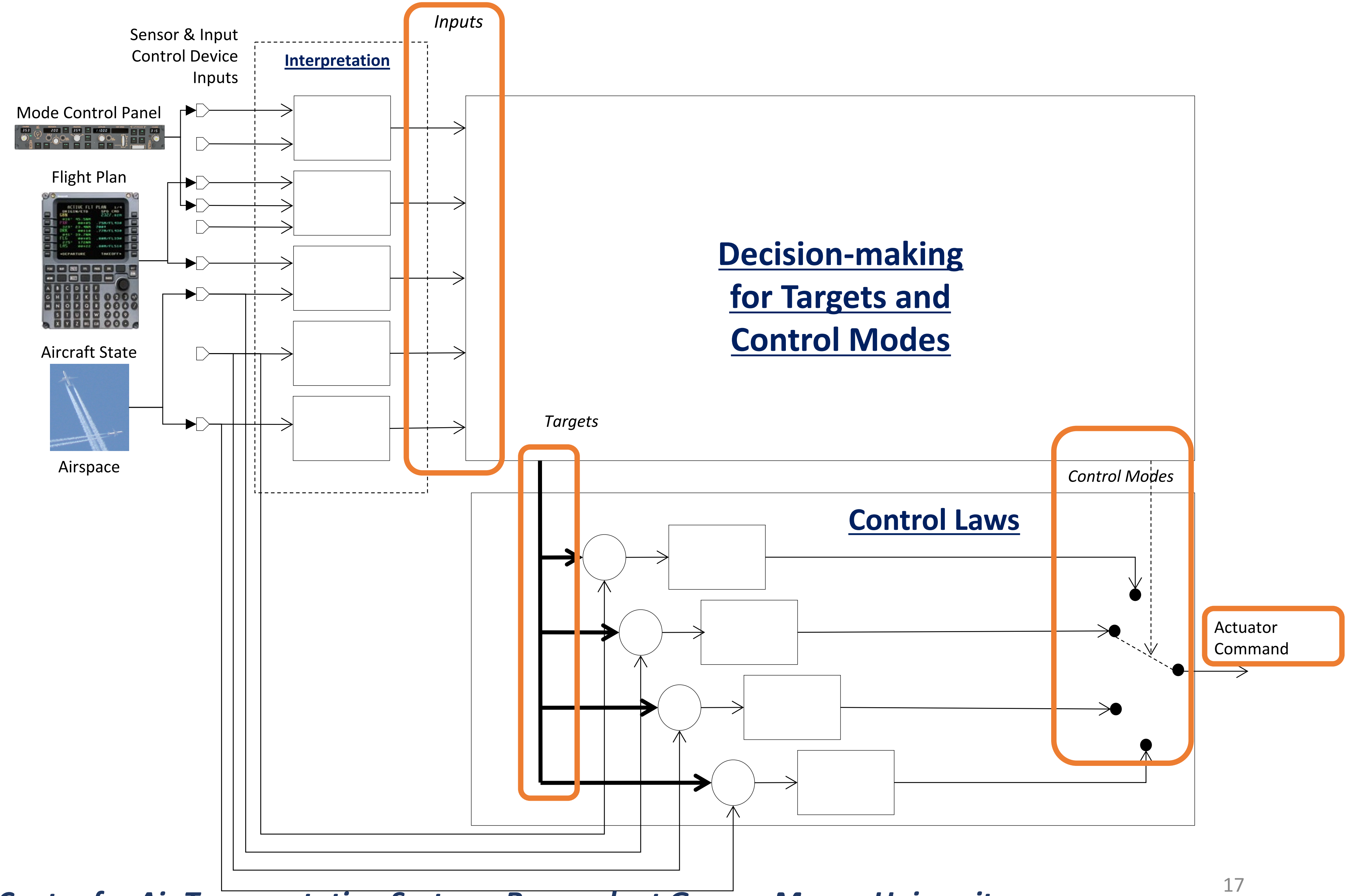
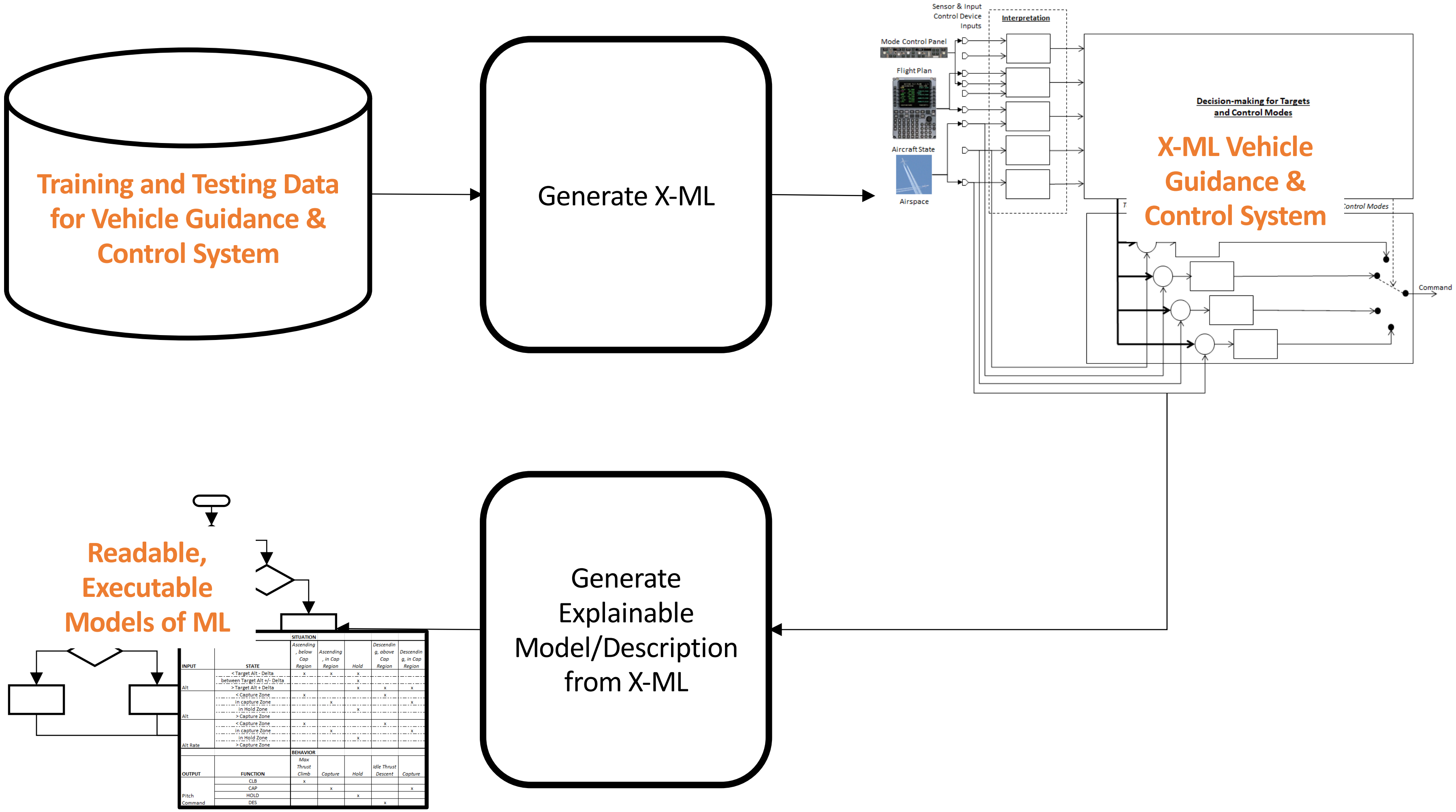


Table of Contents

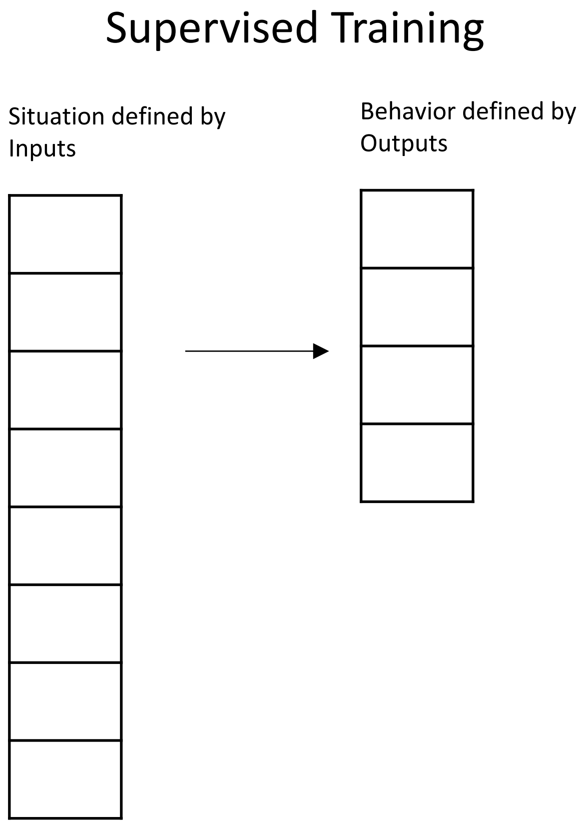
Explainable- Machine Learning (X-ML) for OECS



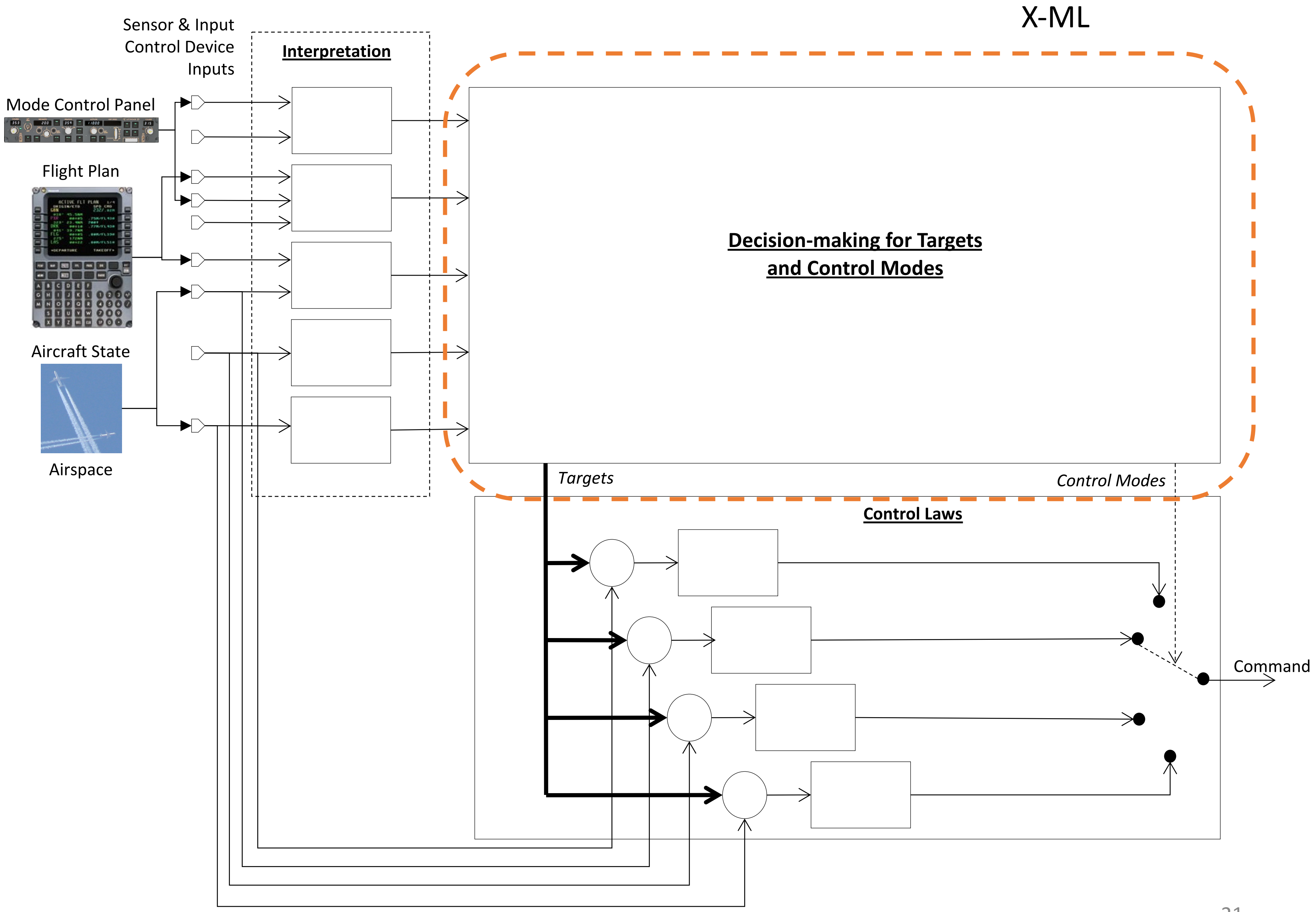
Explainable- Machine Learning (X-ML) for OECS

- Situations = combination of Input States
- Behavior = combination of Output Functions

- X-ML maps Situations to Behaviors
  - Supervised Learning



X-ML is being used for Decision-making for Targets and Control Modes



X-ML Design of Op Embedded Systems

- Steps for X-ML Design of Decision-making for Targets and Controllers
  1. Collect and Process Data from the data-bus
     - Manual control or Automated control operations
     - Manage data for rare/low-frequency events
  2. Supervisory Training/Testing
     - Accuracy/Recall/Precision
     - Rare-events
  3. Simulator/Vehicle Testing
  4. Deployment

X-ML Design of Op Embedded Systems

- Significant reduction in Development Life-cycle
  - 2-3 years - traditional engineering process
  - 2-3 week – X-ML engineering process

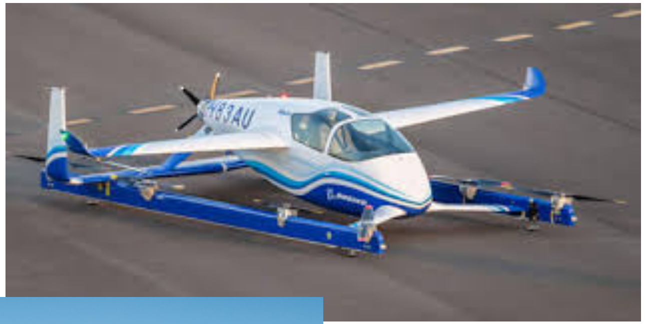
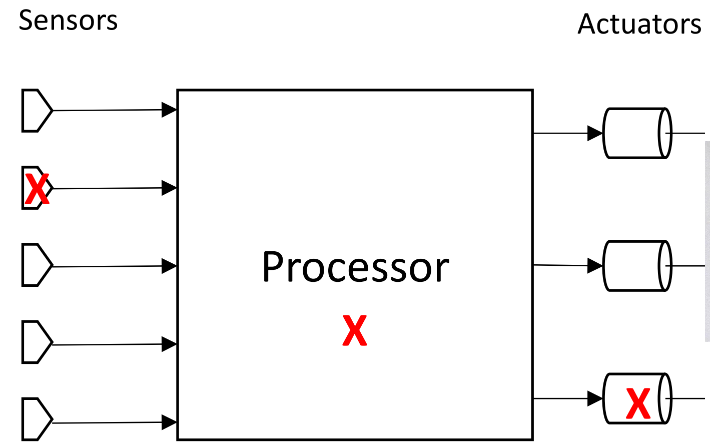


Table of Contents

OECS Accident (Probable) Causes

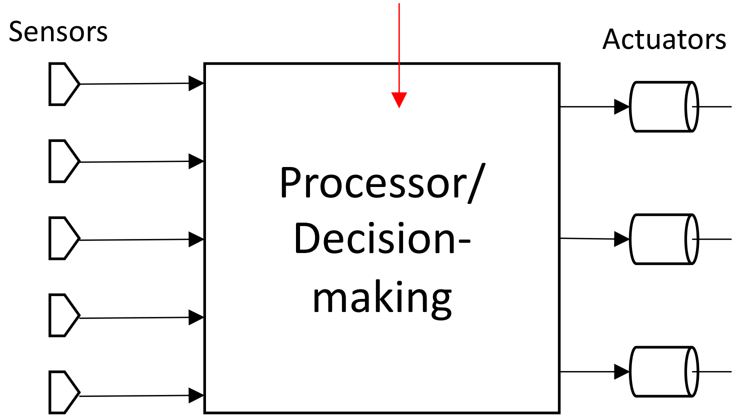
- Equipment Failed
  - Sensor failed
  - Processor failed (e.g. power supply, cable)
  - Actuator failed
  - Mechanical component broke/stuck



- NO Equipment Failed
  - Controlled Flight into Terrain
  - Controlled Flight into Stall
  - Emergent Scenario Accidents /"Normal Accident"

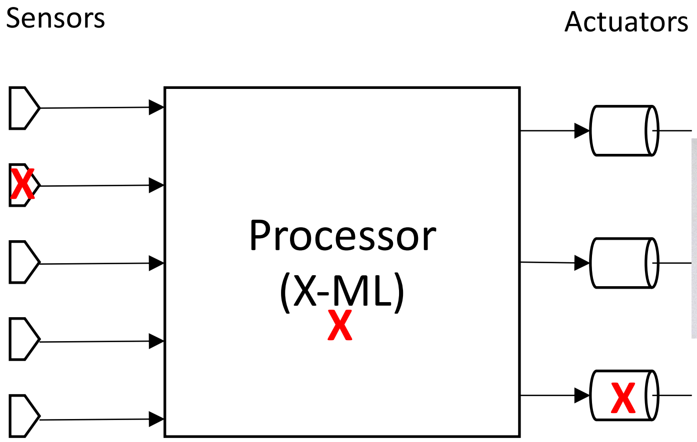
NO EQUIPMENT FAILED MALFUNCTIONS (NEFM)

DESIGN ERROR
(Failure to perform Safe Operations when all equipment is functioning)



X-ML System Failures?

- Equipment Failed
  - Sensor failed
  - Processor failed
  - Actuator failed
  - Mechanical component broke/stuck



- NO Equipment Failed
  - Controlled Flight into Terrain
  - Controlled Flight into Stall
  - Emergent Scenario Accidents/"Normal Accident"

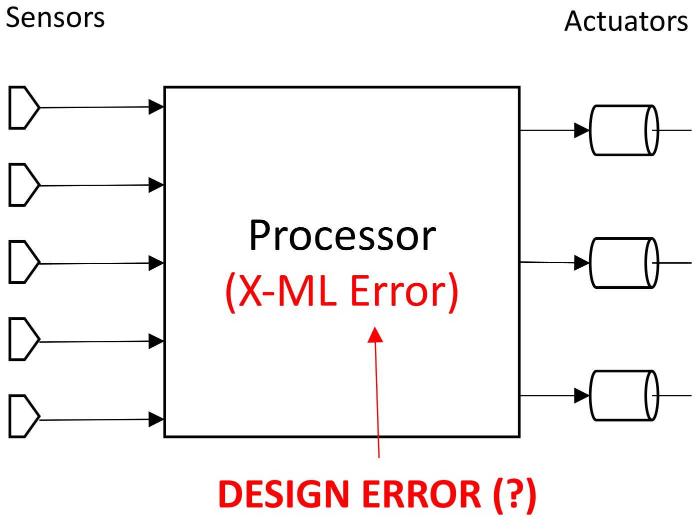
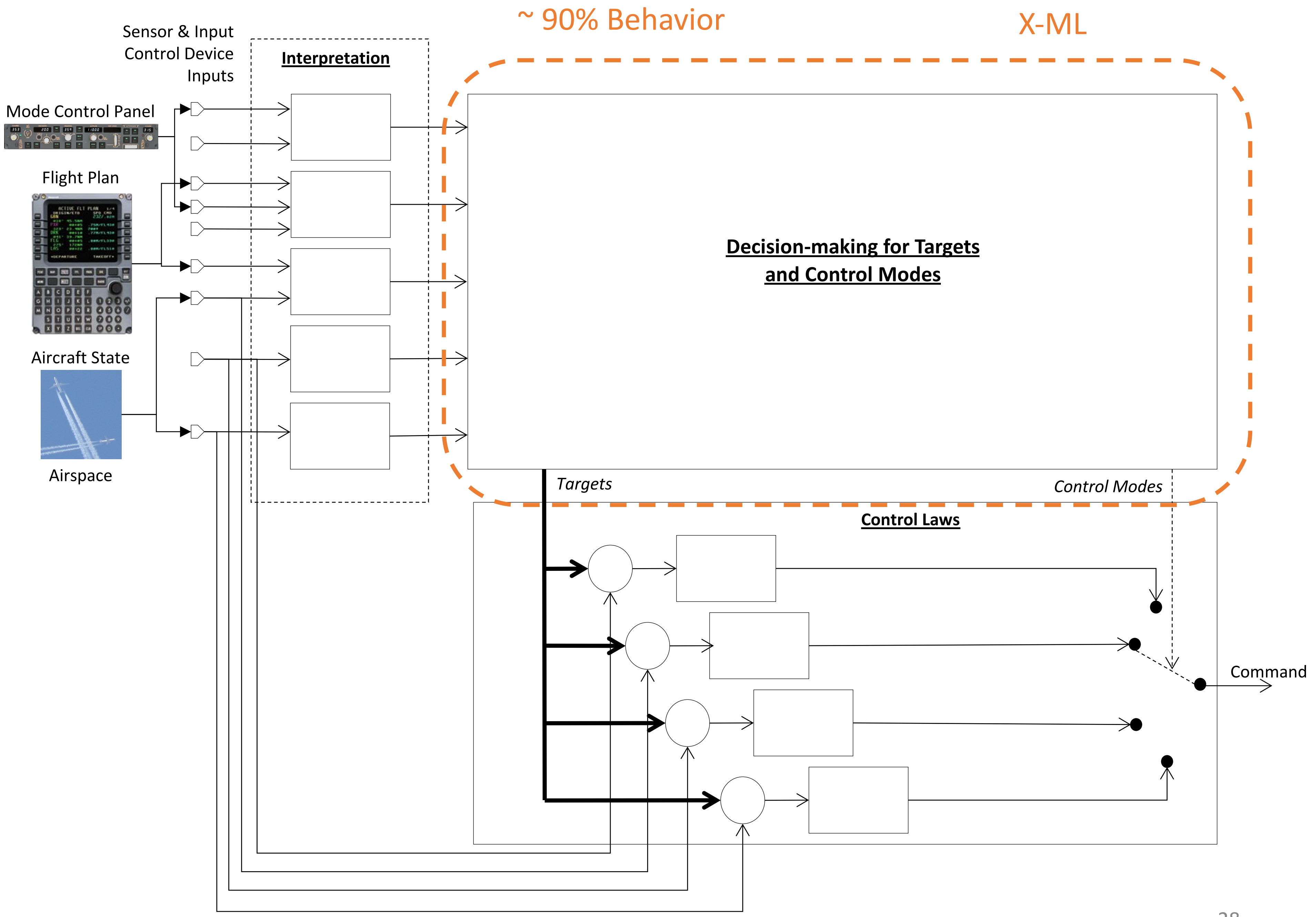
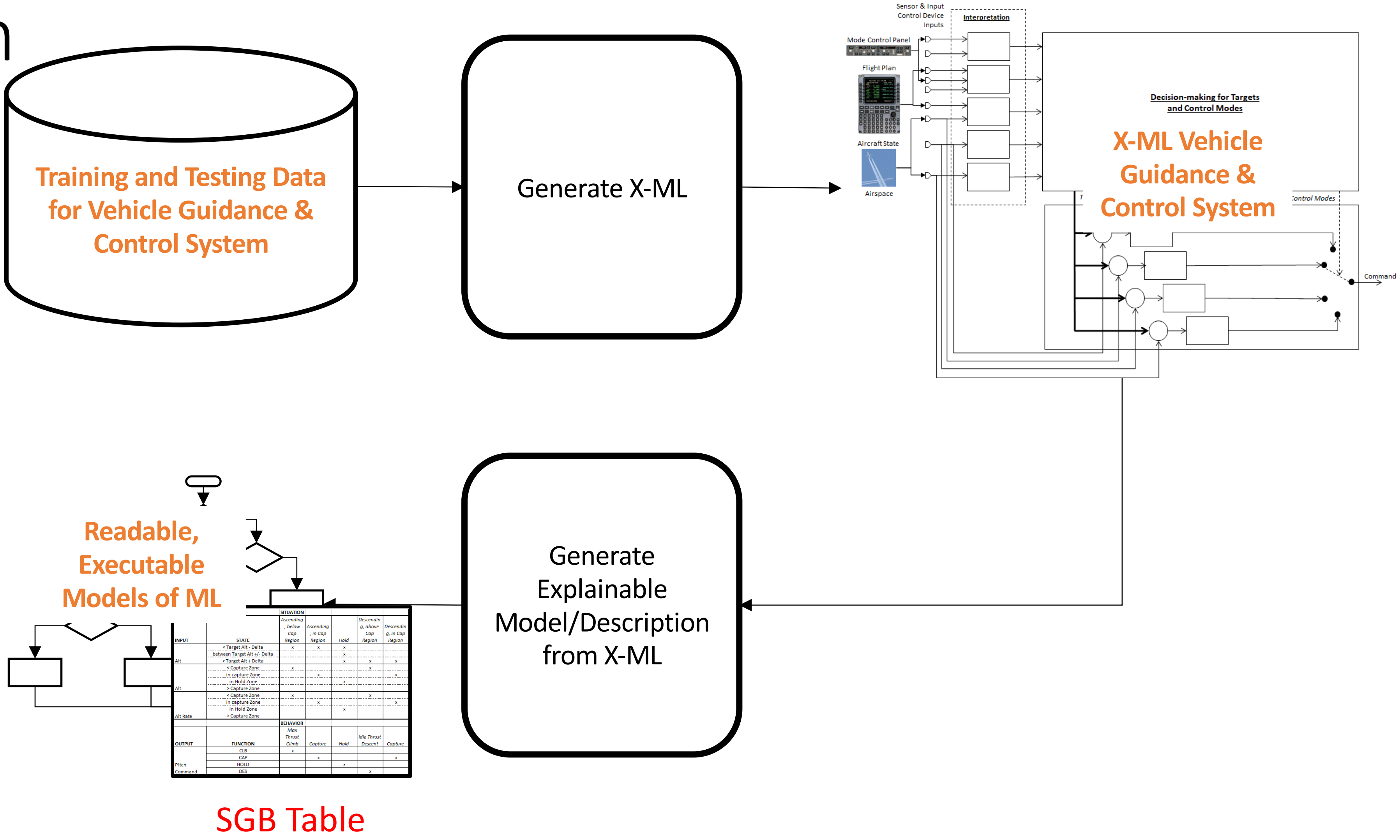


Table of Contents

X-ML is being used for <u>Decision-making</u> for Targets and Control Modes



Trends in Design of Operationally Embedded Systems –Explainable- Machine Learning (X-ML) Design



OECS X-ML Behavior can be modeled by a **Situation-Goal-Behavior Model**

- Operational description
- Executable
- Analyzable

Situation = combination of Input States

Behavior = Selected Targets and Controllers

Situation – Goal – Behavior (SGB) Table

Goals		Airmass	Descent	Late	Descent	Descent Path	Overspeed
Inputs	Situations/ Input States	Aircraft is Descending (without both Prof and FMS Speeds)	Aircraft is descendin g early of D/A Path and Prof/FMS speed engaged	A/C is level late of the D/A Path level at the ref. Alt and the ref. alt	Aircraft is descending late of D/A Path and Prof/FMS speed engaged	Aircraft exceeds speed tolerance while descending on D/A path	Aircraft is level with a speed that exceeds the speed tolerance when ref. Alt is lowered and a/c captures D/A path
VG Type	VNAV /Prof			1	1	1	1
Altitude	Airmass – VNAV/Prof	1	1				
	Airmass - AFS						
Aircraft Altitude	Above distance Referenced D/A path			1	1		
	below distance Referenced D/A path						
Aircraft Speed	Overspeed for D/A path					1	1
	Within speed tolerance for D/A path	1	1	1	1		
Aircraft Altitude	Within D/A Path capture region						
	Not Within D/A Path capture region	1	1	1	1		
Reference Altitude	Has not changed						
	Has changed		1	1			1
Behaviors		Airmass Descent to the D/A path D/A path speed	Referenced recapture using the descent profile	Airmass Descent the D/A the late profile	Referenced to recapture path using descent speed	Airmass Descent D/A path path descent	Referenced around the at the D/A speed profile
Altitude Target	M:Climb/Cruise						
	M:Descent/Approach	Descent/ Altitude	Approach Target	Descent/ Altitude	ApproachTarget	Descent/ Altitude	ApproachTarget
Speed Target	M:Late descent			Late Speed	Descent Target		
	M: Descent/Approach					Descent/ Speed	Approach Target
	M: Airmass Descent	Airmass Speed	Descent Target				
	P: engine-out						
Speed/	P: THRUST |HOLD						

OECS: Design Error Archetypes

1. SGB Table Missing Input
   - Design is ***absent one or more of the required inputs*** (i.e. sensors/data feeds) to identify one or more of the operational situations that must be covered by the operationally embedded system
2. SGB Table Missing Input/State Combinations
   - Given all the required inputs, the design is ***absent one or more combinations of input states*** to respond to *all* the operational situations that must be covered by the operationally embedded system
3. SGB Table Missing Mapping between Input/State Combinations to Behaviors
   - Given the required inputs to support all the combinations of input states and all the combinations of input states, the design is ***absent one or more the correct mappings*** between operational situations and appropriate behaviors

Design Error Archetype #1

1 - Missing Input

- Design is ***absent one or more of the required inputs*** (i.e. sensors/data feeds) to identify one or more of the operational situations that must be covered by the operationally embedded system

Missing an Input

Goals		Airmass	Descent	Late	Descent	Descent Path	Overspeed
Inputs	Situations/ Input States	Aircraft is Descending (without both Prof and FMS Speeds)	Aircraft is descendin g early of D/A Path and Prof/FMS speed engaged	A/C is level late of the D/A Path level at the ref. Alt and the ref. alt	Aircraft is descending late of D/A Path and Prof/FMS speed engaged	Aircraft exceeds speed tolerance while descending on D/A path	Aircraft is level with a speed that exceeds the speed tolerance when ref. Alt is lowered and a/c captures D/A path
VG Type Altitude	VNAV /Prof			1	1	1	1
	Airmass – VNAV/Prof	1	1				
	Airmass - AFS						
Aircraft Altitude	Above distance Referenced D/A path			1	1		
	below distance Referenced D/A path						
Aircraft Speed	Overspeed for D/A path					1	1
	Within speed tolerance for D/A path	1	1	1	1		
Aircraft Altitude	Within D/A Path capture region						
	Not Within D/A Path capture region	1	1	1	1		
Reference Altitude	Has not changed						
	Has changed		1	1			1
Behaviors		Airmass Descent to the D/A path D/A path speed	Referenced recapture using the descent profile	Airmass Descent the D/A the late profile	Referenced to recapture path using descent speed	Airmass Descent D/A path path descent	Referenced around the at the D/A speed profile
Altitude Target	M:Climb/Cruise						
	M:Descent/Approach	Descent/ Altitude	Approach Target	Descent/ Altitude	ApproachTarget	Descent/ Altitude	ApproachTarget
Speed	M:Late descent			Late Speed	Descent Target		
Target	M: Descent/Approach					Descent/ Speed	Approach Target
	M: Airmass Descent	Airmass Speed	Descent Target				
	P: engine-out						
Speed/	P: THRUST |HOLD						

“Say … whats a mountain goat doing up here?”

Design Error Archetype #1

1 - Missing Input

- Design is ***absent one or more of the required inputs*** (i.e. sensors/data feeds) to identify one or more of the operational situations that must be covered by the operationally embedded system
  - **Windshear Alerting and Guidance Mandate**
    - Aircraft automation/flight-crews did not distinguish between Windshear conditions and high wind
      - Windshear – headwind transitions (almost instantaneously) to tailwind
  - **Traffic Collision Avoidance Mandate**
    - Aircraft automation/flight-crews did not have information about near-term collision trajectories

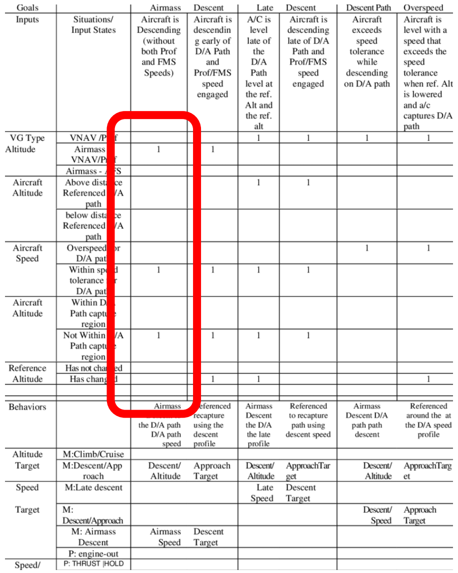
“Say … whats a mountain goat doing up here?”

Design Error Archetype #2

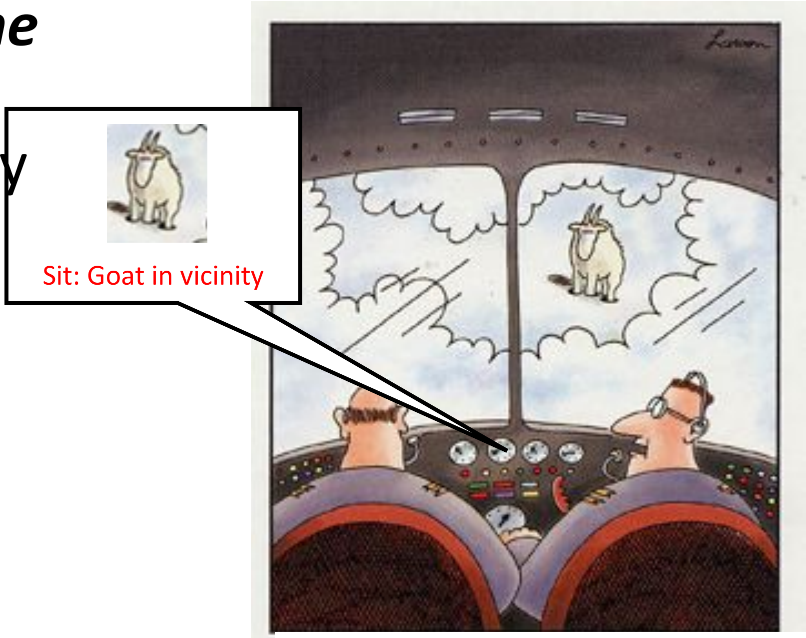
2 - Missing Input/State Combinations

- Given all the required inputs, the design is ***absent one or more combinations of input states*** to respond to *all* the operational situations that must be covered by the operationally embedded system

Missing Situation (i.e. combination of Input States)



Goals / Inputs	Situations/ Input States	Airmass: Aircraft is Descending (without both Prof and FMS Speeds)	Descent: Aircraft is descending early of D/A Path and Prof/FMS speed engaged	Late: A/C is level late of the D/A Path level at the ref. Alt and the ref. alt	Descent: Aircraft is descending late of D/A Path and Prof/FMS speed engaged	Descent Path: Aircraft exceeds speed tolerance while descending on D/A path	Overspeed: Aircraft is level with a speed that exceeds the speed tolerance when ref. Alt is lowered and a/c captures D/A path
VG Type Altitude	VNAV /P f			1	1	1	1
	Airmass VNAV/P f	1	1				
	Airmass - S						
Aircraft Altitude	Above dist ce Referenced /A path			1	1		
	below dist ce Referenced /A path						
Aircraft Speed	Overspeed r D/A pat					1	1
	Within sp d tolerance r D/A pat	1	1	1	1		
Aircraft Altitude	Within D Path capt e region						
	Not Within /A Path capt e region	1	1	1	1		
Reference Altitude	Has not cha ed						
	Has chang d		1	1			1
Behaviors		Airmass the D/A path D/A path speed	eferenced recapture using the descent profile	Airmass Descent the D/A the late profile	Referenced to recapture path using descent speed	Airmass Descent D/A path path descent	Referenced around the at the D/A speed profile
Altitude Target	M:Climb/Cruise						
	M:Descent/App roach	Descent/ Altitude	Approach Target	Descent/ Altitude	ApproachTar get	Descent/ Altitude	ApproachTarg et
Speed	M:Late descent			Late Speed	Descent Target		
Target	M: Descent/Approach					Descent/ Speed	Approach Target
	M: Airmass Descent	Airmass Speed	Descent Target				
	P: engine-out						
Speed/	P: THRUST |HOLD						



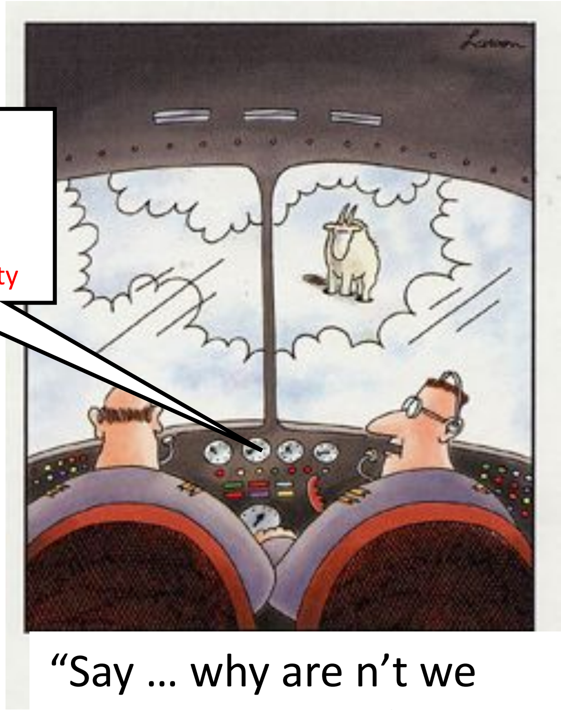
"Say … why are n't we turning to avoid the mountain goat"

Design Error Archetype #2

2 - Missing Input/State Combinations

- Given all the required inputs, the design is ***absent one or more combinations of input states*** to respond to *all* the operational situations that must be covered by the operationally embedded system
  - **Las Vegas Autonomous Shuttle Bus Accident**
    - **Automation did not resolve situation of Tractor Trailer crossing street vs. Tractor Trailer backing-up into perpendicular alley**

      Sherry, et. al. (2020) Autonomous Systems Design, Testing, and Deployment: Lessons Learned from the Deployment of an Autonomous Shuttle Bus
  - **Air France 447 Accident**
    - **Automation did not know how to handle situation of discrepancy in airspeed from triple redundant airspeed sensor data**
  - **Turkish Airlines 1951**
    - **Automation did not resolve situation of discrepancy between Radar Altimeter and Barometric Pressure Altitude**



"Say … why are n't we turning to avoid the mountain goat"

Design Error Archetype #3

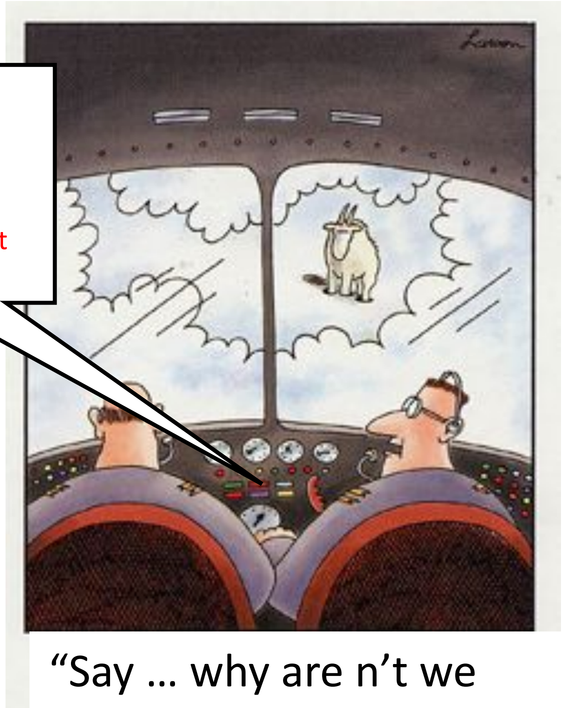
3 - Missing Mapping between Input/State Combinations to Behaviors

- Given the required inputs *and* all the combinations of input states, the design is ***absent one or more the correct mappings*** between operational situations and appropriate behaviors

Goals		Airmass	Descent	Late	Descent	Descent Path	Overspeed
Inputs	Situations/ Input States	Aircraft is Descending (without both Prof and FMS Speeds)	Aircraft is descendin g early of D/A Path and Prof/FMS speed engaged	A/C is level late of the D/A Path level at the ref. Alt and the ref. alt	Aircraft is descending late of D/A Path and Prof/FMS speed engaged	Aircraft exceeds speed tolerance while descending on D/A path	Aircraft is level with a speed that exceeds the speed tolerance when ref. Alt is lowered and a/c captures D/A path
VG Type Altitude	VNAV /Prof			1	1	1	1
	Airmass – VNAV/Prof	1	1				
	Airmass - AFS						
Aircraft Altitude	Above distance Referenced D/A path			1	1		
	below distance Referenced D/A path						
Aircraft Speed	Overspeed for D/A path					1	1
	Within speed tolerance for D/A path	1	1	1	1		
Aircraft Altitude	Within D/A Path capture region						
	Not Within D/A Path capture region	1	1	1	1		
Reference Altitude	Has not changed						
	Has changed		1	1			1
Behaviors		Airmass Descent to the D/A path D/A path speed	Referenced recapture using the descent profile	Airmass Descent the D/A the late profile	Referenced to recapture path using descent speed	Airmass Descent D/A path path descent	Referenced around the at the D/A speed profile
Altitude Target	M:Climb/Cruise						
	M:Descent/App roach	Descent/ Altitude	Approach Target	Descent/ Altitude	ApproachTar get	Descent/ Altitude	ApproachTarg et
Speed	M:Late descent			Late Speed	Descent Target		
Target	M: Descent/Approach					Descent/ Speed	Approach Target
	M: Airmass Descent	Airmass Speed	Descent Target				
	P: engine-out						
Speed/	P: THRUST |HOLD						

Missing or Incorrect Mapping of Situation to Behavior

Sit: Goat straight ahead

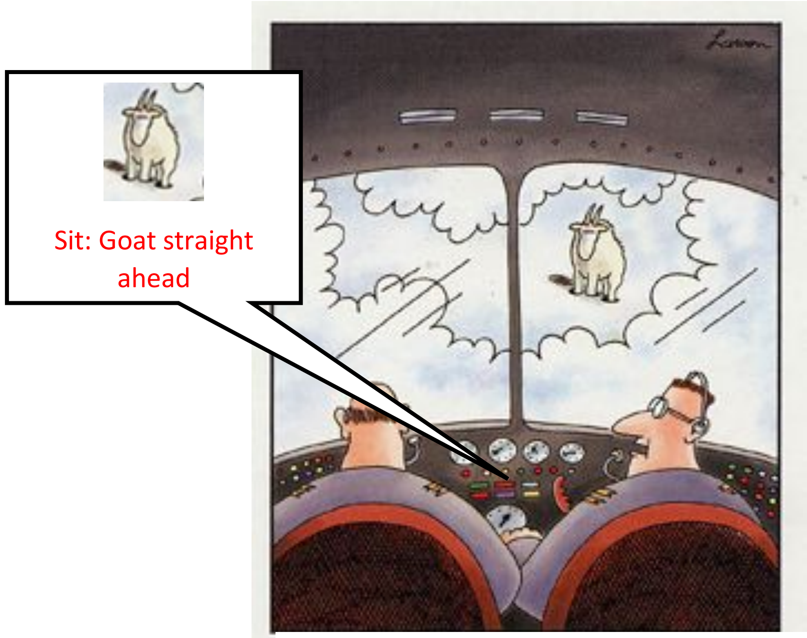


"Say … why are n't we turning to avoid the mountain goat"

Design Error Archetype #3

3 - Missing Mapping between Input/State Combinations to Behaviors

- Given the required inputs to support all the combinations of input states and all the combinations of input states, the design is ***absent one or more the correct mappings*** between operational situations and appropriate behaviors
  - **Asiana Air 241 Accident**
    - **"Human/Automation" System did not respond to under-speed condition**



"Say … why are n't we turning to avoid the mountain goat"

Challenges for Design X-ML Op Embedded Systems

X-ML Design is only as good as the completeness of the training/testing data

1. Training/Testing data is missing inputs

2. Training/Testing data has all the input variables, *but* Training/Testing data is missing combinations of Inputs/States

3. Training/Testing data has all the input variables, *and all* combinations of Inputs/States, *but* Training/Testing data is missing scenarios that map input/state combinations to appropriate output behaviors

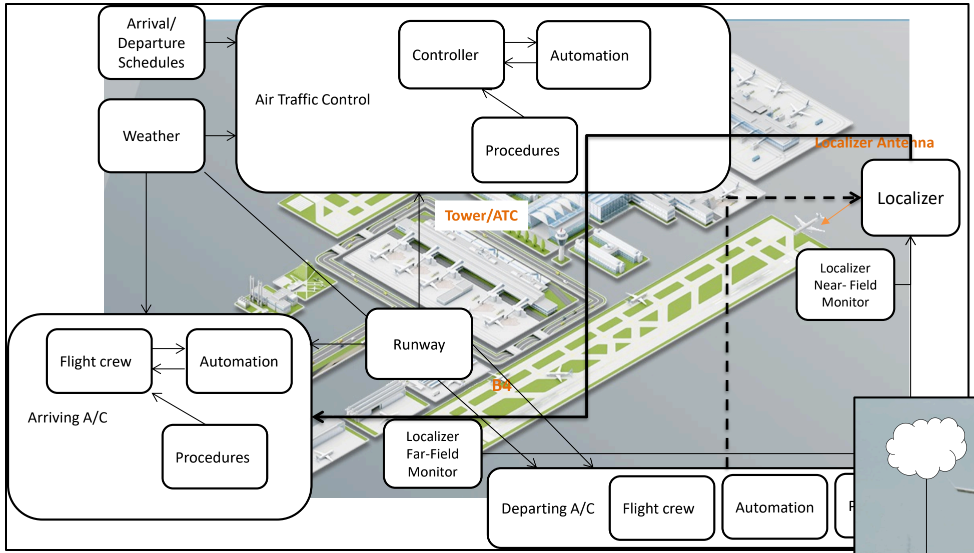
Table of Contents

Mitigating Issues with X-ML Op Embedded Systems

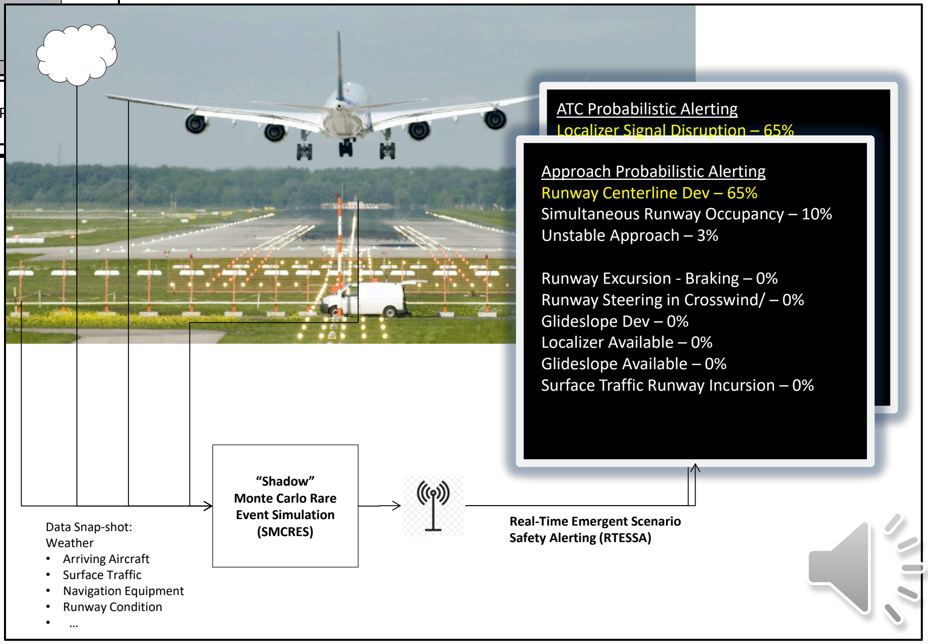
1. Training/Testing data is missing inputs
   - SME review Situation/Behaviors
   - Scenario Analysis/Hazard Analysis
   - Fast Time Emergent Scenario Simulation (FTESS)
2. Training/Testing data has all the input variables, *but* Training/Testing data is missing combinations of Inputs/States
   - Check all combinations of Input/States are included
     - SGB Tables provides a quick/easy way to check for completeness
3. Training/Testing data has all the input variables, *and all* combinations of Inputs/States, *but* Training/Testing data is missing scenarios that map input/state combinations to appropriate output behaviors
   - Check every Situation is mapped to a Behavior
     - SGB Tables provides a quick/easy way to check for mapping
   - SME Review Behaviors for each combination of Inputs/States with SME

Fast-Time Emergent Scenario Simulation (FTESS)

- Finding situations not in the design before they occur
- Situations are interactions between system-of-system components
- Run simulation 365/24/7 even after the system is "certified"/fielded

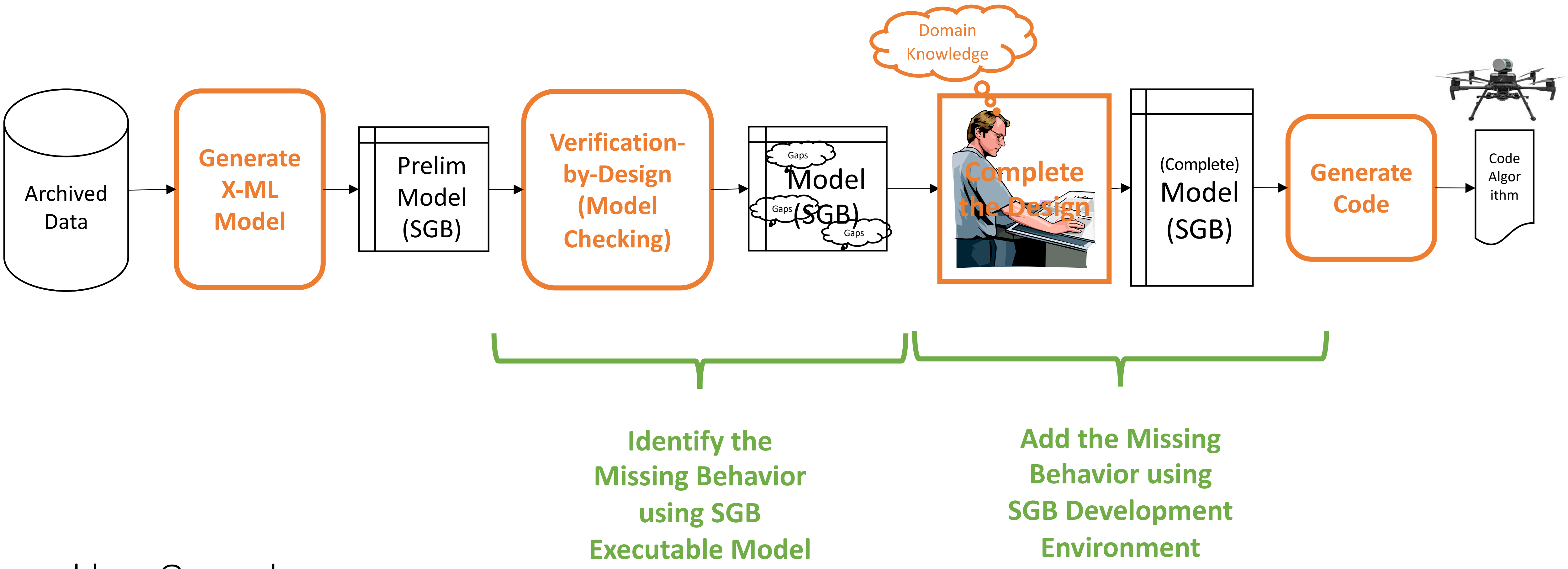


Interaction between System-of-System components



Run in Shadow-Mode even after Certification Fielding

Collaborative Functional Design Using X-ML



lsherry@gmu.edu

Table of Contents

Conclusion

- Using X-ML for Operationally Embedded Control Systems:
  - has tremendous potential
  - requires mediation to account for mission situations not in the data
    1. Missing Inputs
    2. Missing Combination of Input/States
    3. Missing mapping of combination of Input/States and Behaviors
- There are no “short-cuts” to designing complex systems
  - X-ML Designs can only be based on data set provided:
    - Situations-Behavior Pairs
- X-ML does provide a means to reduce development time

lsherry@gmu.edu