

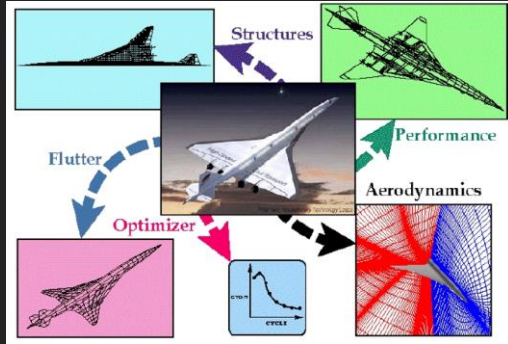
Mod-Sim and Decision-Making in Aerospace Domain

Natalia M. Alexandrov
NASA Langley Research Center

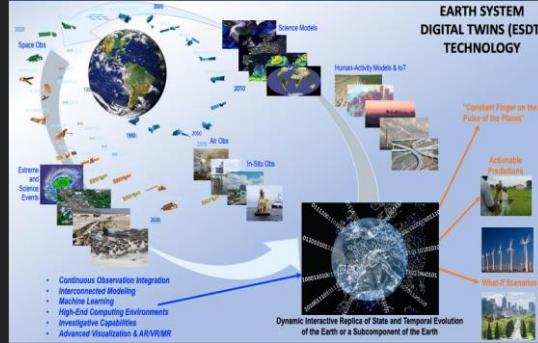
Assessing the Reliability of Complex, Dynamic Modeling and Simulation

NASEM, 17 June 2024

Computational modeling & simulation of engineering and natural systems advancing for decades



Vehicle Systems



Earth Science Systems



Aviation Systems



Space Technology Systems

- Analysis and design-oriented variable-fidelity models (usually differential equation solvers)
- Single-discipline and multidisciplinary design optimization under uncertainty (robust, reliability, risk-based)
- Control
- Inverse problems

- High complexity
- Traditionally, differential equations
- Increased use of machine learning
- Current target: Earth System Digital Twins (ESDT) Technology

- High-complexity human-machine teams
- Long experience developing advanced airspace capabilities that rely on simulations and field integration
- Advanced Testbed integrates facilities and live flightdeck with simulated air traffic and advanced control concepts

- High complexity
- Operation in extreme environments
- High degree of uncertainty
- Human-machine teaming
- Machine autonomy
- Digital twin

Consider Trust and Trustworthiness in a Use Case



How to achieve **trustworthiness and justified trust** in autonomous cyber-physical-human (CPH) systems in **safety-critical, time-critical, and certification-dependent** environments?



ATTRACTOR
TEAM

FIN

06/12/2020

Step 2 Feasibility Review



FY20

FY19

FY18

...



The ATTRACTOR Challenge

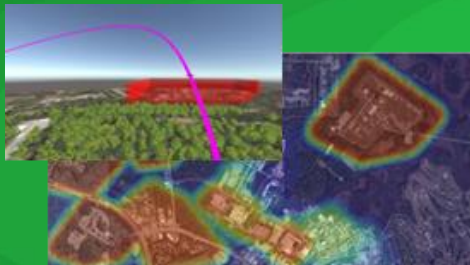
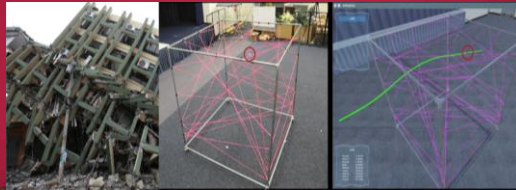


How to achieve **trustworthiness and justified trust** in autonomous cyber-physical-human (CPH) systems in **safety-critical, time-critical, and certification-dependent** environments?

Multidisciplinary Components of
Complex Multi-agent Systems

Measurable **trustworthiness** of
fundamental functional
components

Context:
Trajectories



Sound multi-objective
decision-making

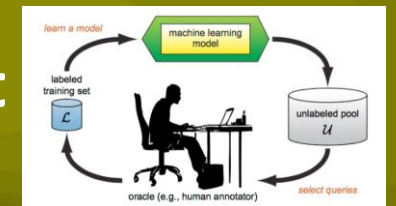
Heterogeneous multi-agent (including human-machine)
teaming is essential



Teaming

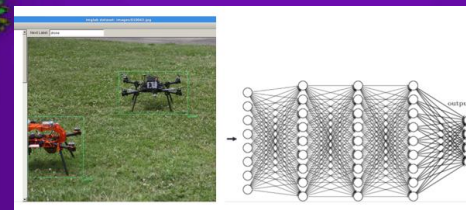
Efficient two-way
learning/ training loop.

**Learning &
Justifiable Trust**



**Good Outcomes:
Basis for Trust**

Explainability

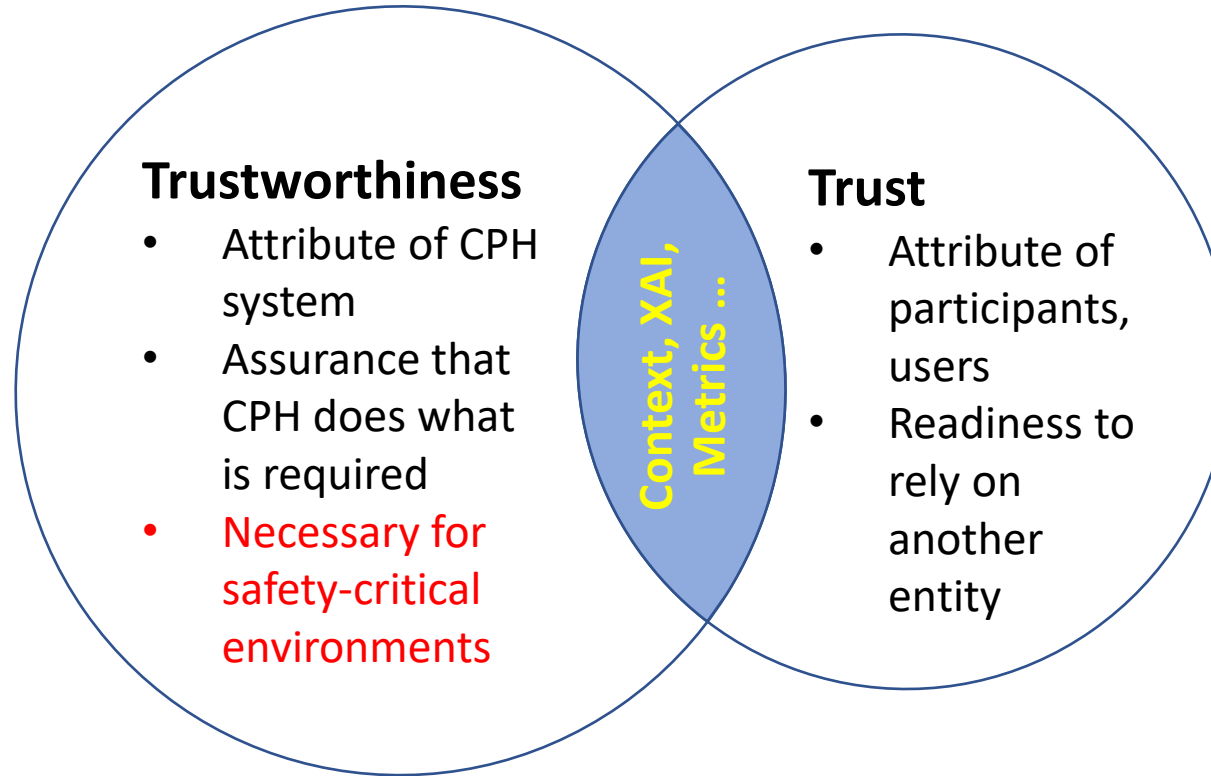


These are two UAVs. Because... They
are quadrotors. They have four
propellers. They have four legs.

Trustworthiness vs. Trust

Some components informing trustworthiness:

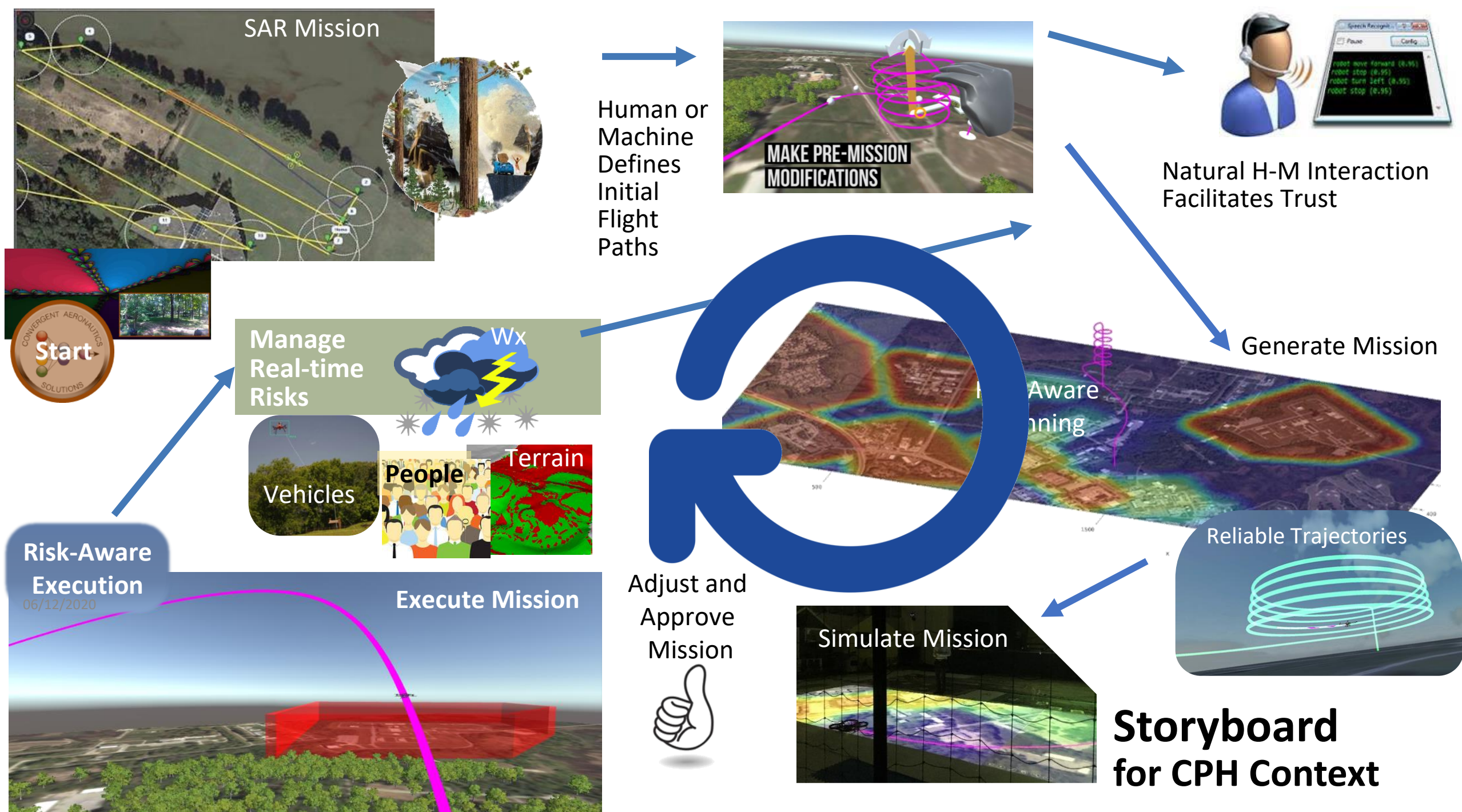
- Context
- Trustworthiness models
- Physics (e.g., trajectory planning)
- Multi-objective decision-making under uncertainty
- Anomaly detection
- V&V (stress testing)
- Persistent modsim
- XAI (for performance)
- Metrics (thresholds)
- Multi-agent system
- Sensors



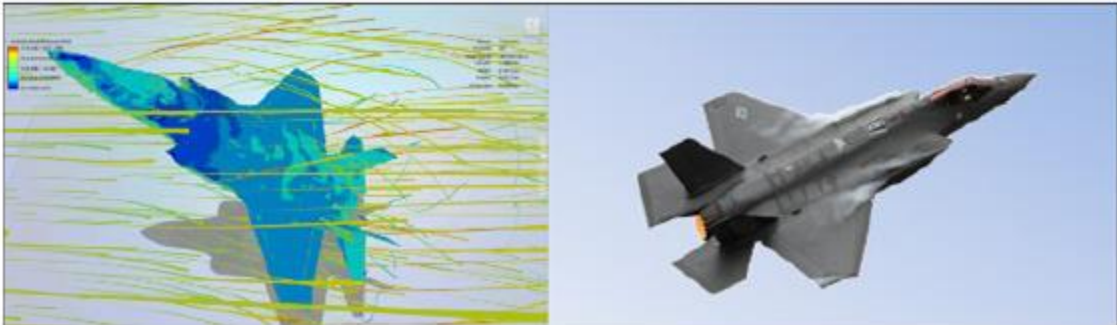
Some components informing trust:

- Context
- Trustworthiness models
- Natural HMI
- Two-way learning
- Interaction history
- XAI (for transparency)
- Metrics (thresholds)

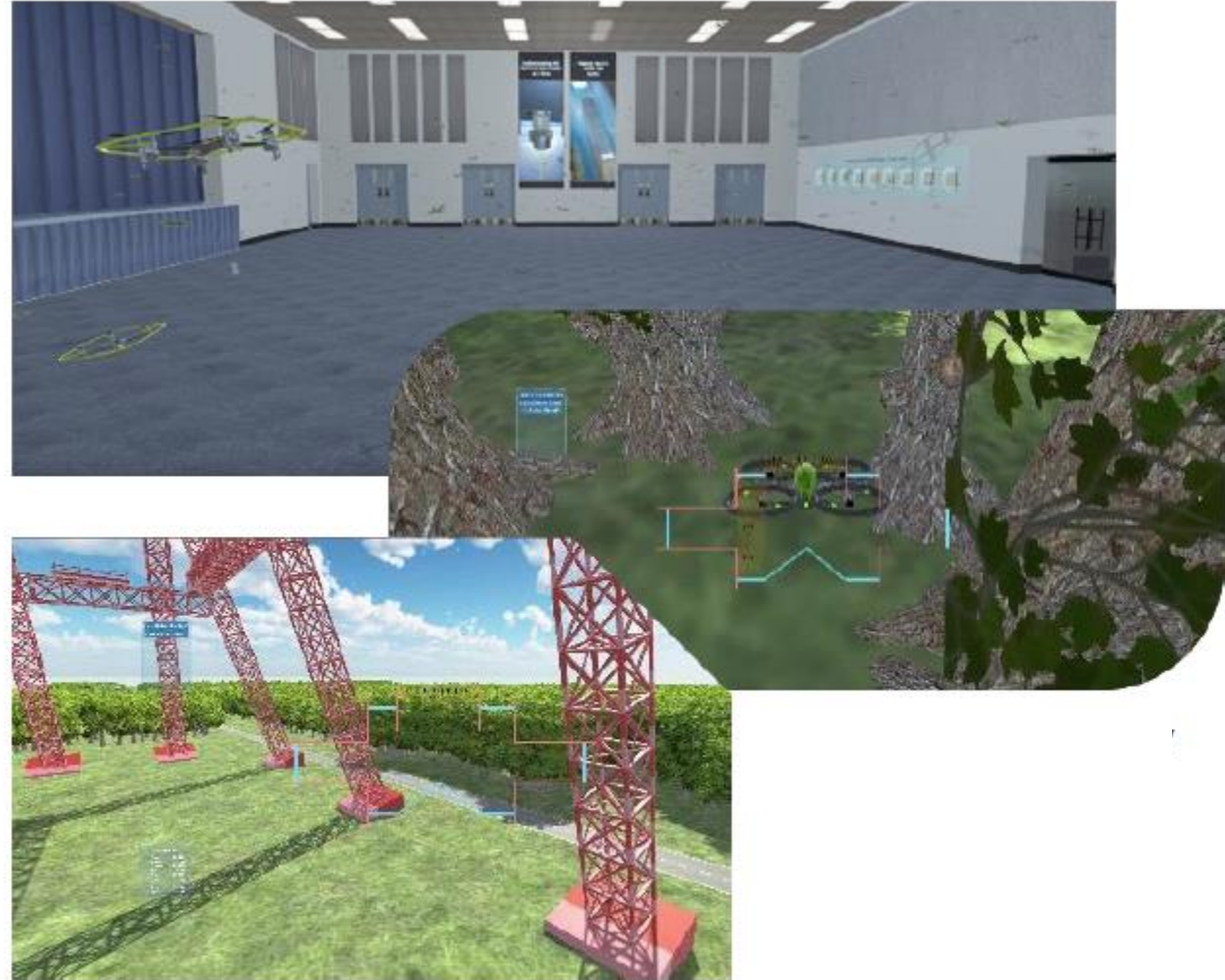
- Certification is a set of functional requirements and bounds that imply trustworthiness.
- To the best of our knowledge, there are currently no certification criteria for autonomous CPH systems.



Essential Component: Persistent Mod-Sim



Digital Twin of a mechanical artifact or an environment



“Digital Twin” of a “living system”

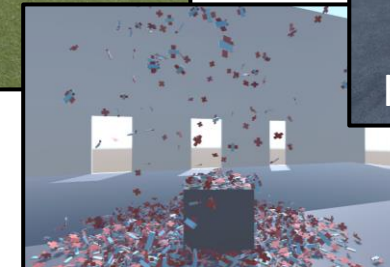
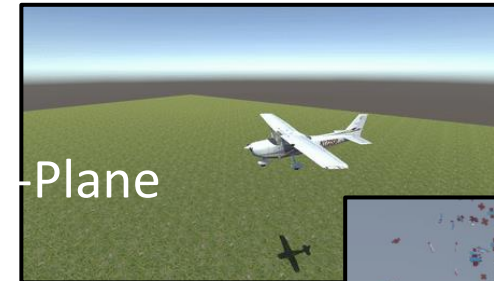
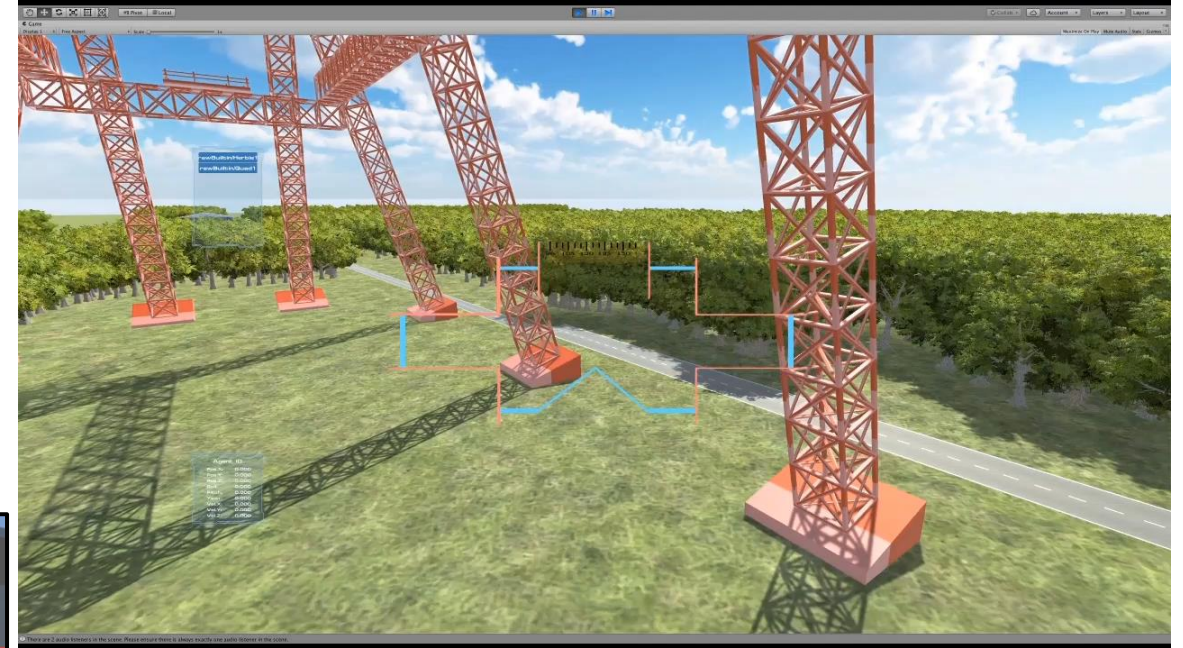
Persistent Mixed-Reality Sim-to-Flight “DT”



Baseline Environment for Autonomous Modeling (BEAM, Ben Kelley)



- Multi-user
- Modular plug-in capability
- No/minimal dependency on central servers
- Extensible: sensors, vehicles, environments, air traffic control, etc.
- Common communication interface (DDS) supports external research code



MMORPG

Unity-based
simulation
environment

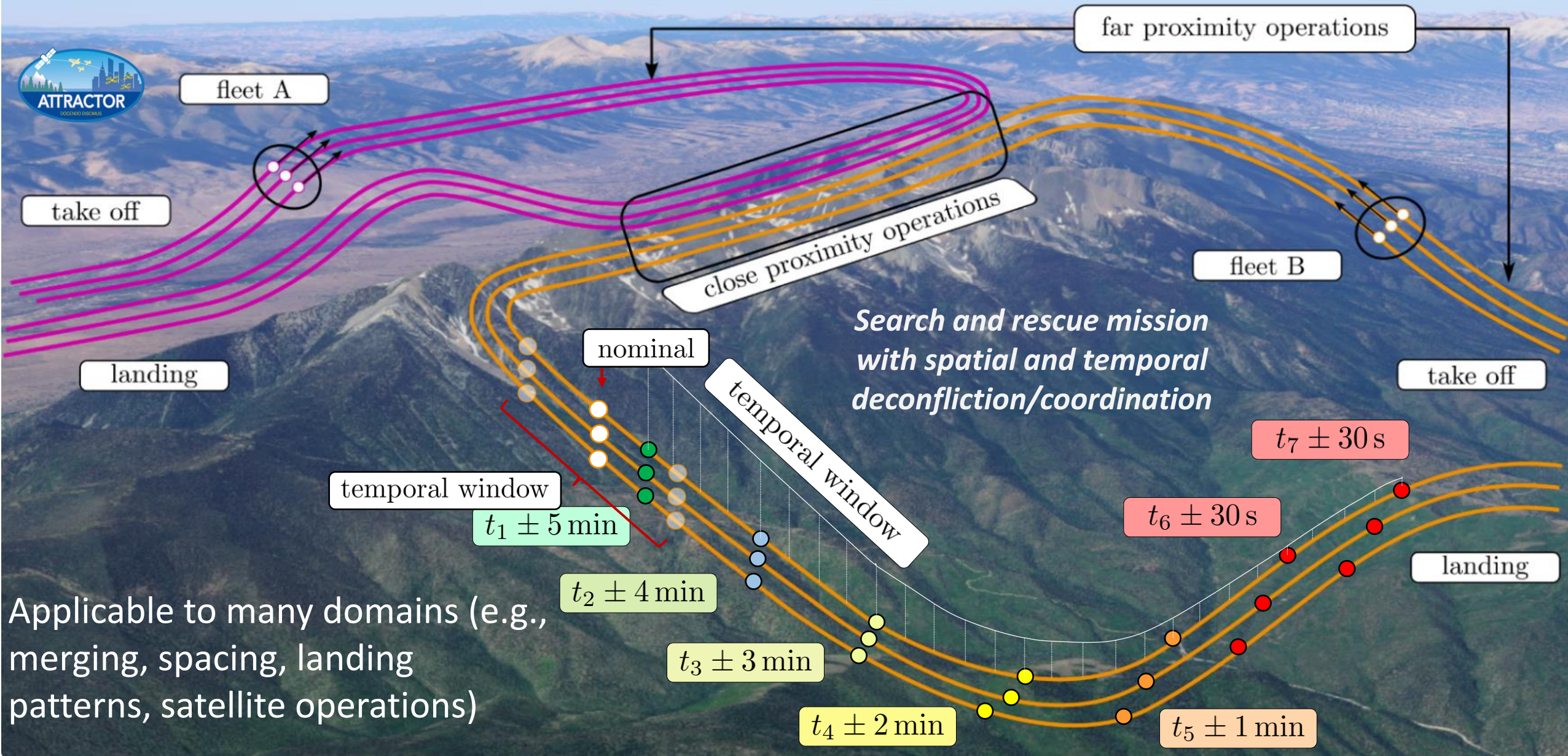
Mixed-Reality Environment (Mathew Vaughan)



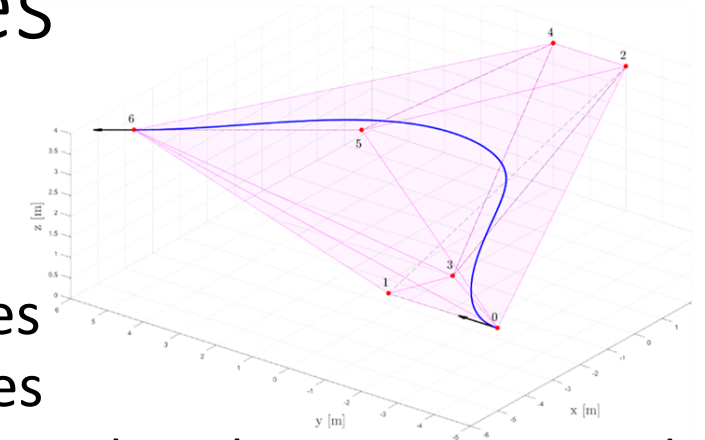
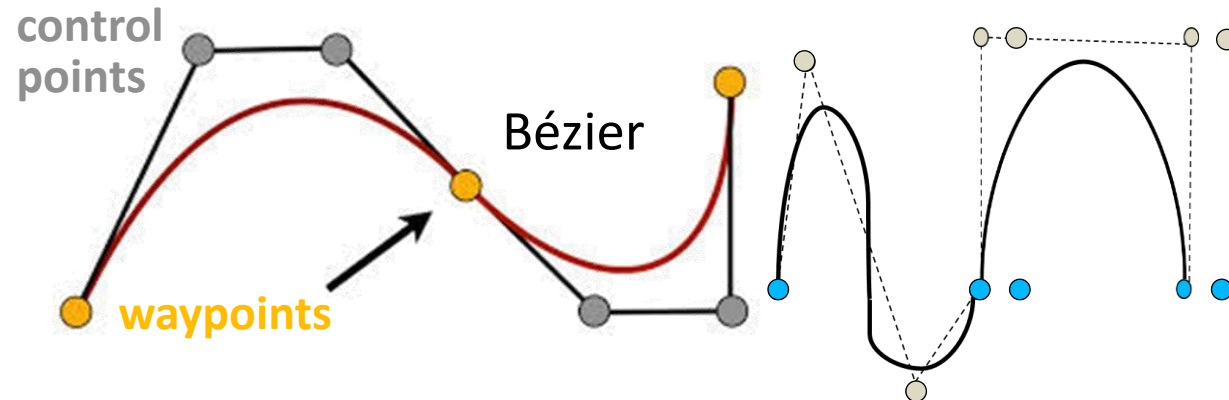
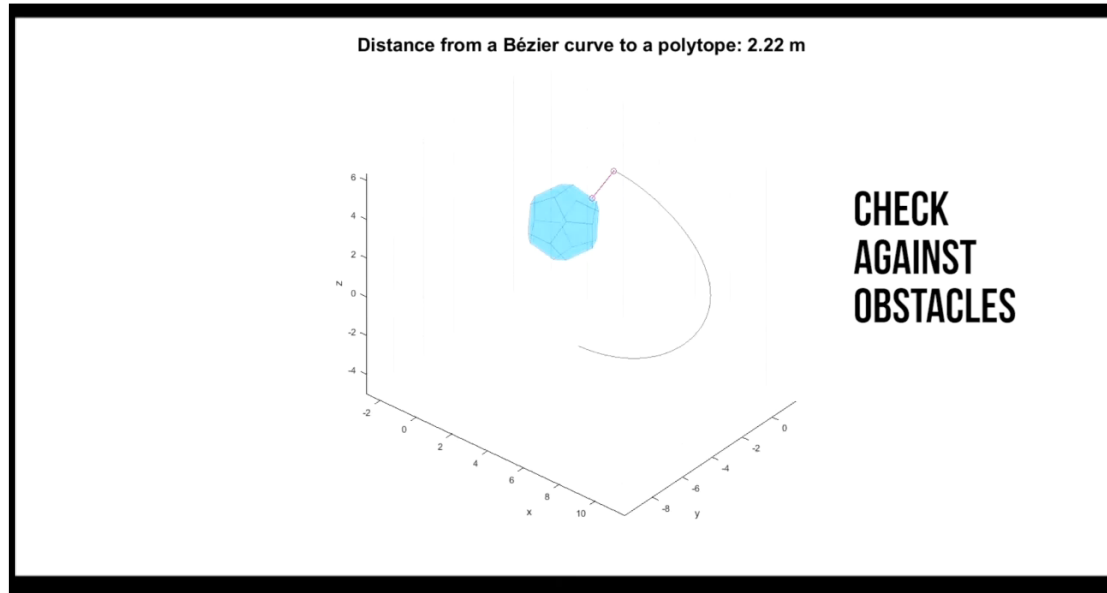
Example of a SAR mission



Under the Hood: Trajectory Planning (Javier Puig-Navarro)



Bezier-Based Trajectories and Polytopes



Components

- Temporal Queries
- Proximity Queries
 - To limit the scope, obstacles are represented as polyhedrons
 - Trajectories are represented by a set of time parameterized polynomial (Bézier) curves

Enables

- Time Coordination/Deconfliction
- Spatial Coordination (Proximity Ops)
- Trajectory-Based Operations

Tolerance Verification = TQ + PQ

Scalable: required for multi-vehicle operations

Recent Developments: Design of Safe Separation Bounds under Bounded Uncertainties

- Consider a fleet of n **heterogeneous UAS** with a set of unique identification numbers

$$\mathcal{I} := \{1, \dots, n\}$$

- Each UAS has been assigned a trajectory

$$p_{d_i} : [t_{ini_i}, t_{end_i}] \rightarrow \mathbb{R}^3,$$

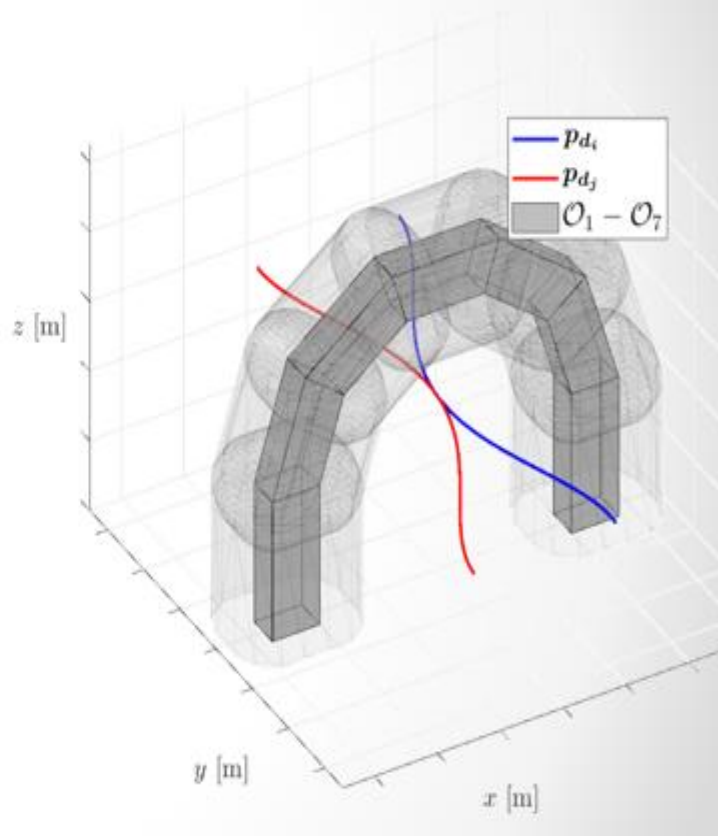
- Trajectories are **temporally deconflicted**

$$\|p_{d_i}(t_d) - p_{d_j}(t_d)\| \geq d_{s_{i,j}}$$

for all $t_d \in [t_{ini_i}, t_{end_i}] \cap [t_{ini_j}, t_{end_j}]$,
as opposed to **spatially deconflicted**

$$\|p_{d_i}(t_{d_i}) - p_{d_j}(t_{d_j})\| \geq d_{s_{i,j}}$$

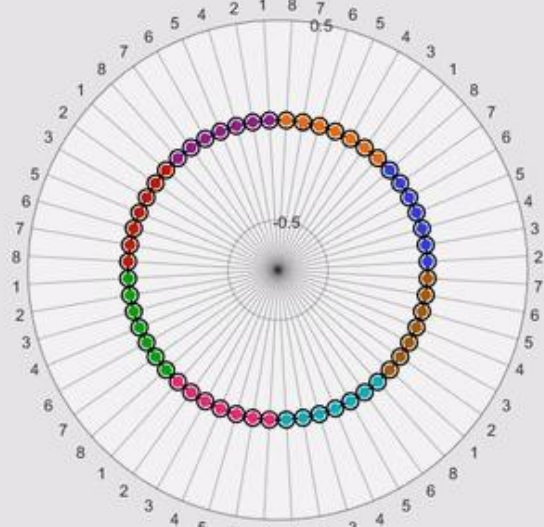
for all $t_{d_i} \in [t_{ini_i}, t_{end_i}]$, $t_{d_j} \in [t_{ini_j}, t_{end_j}]$.



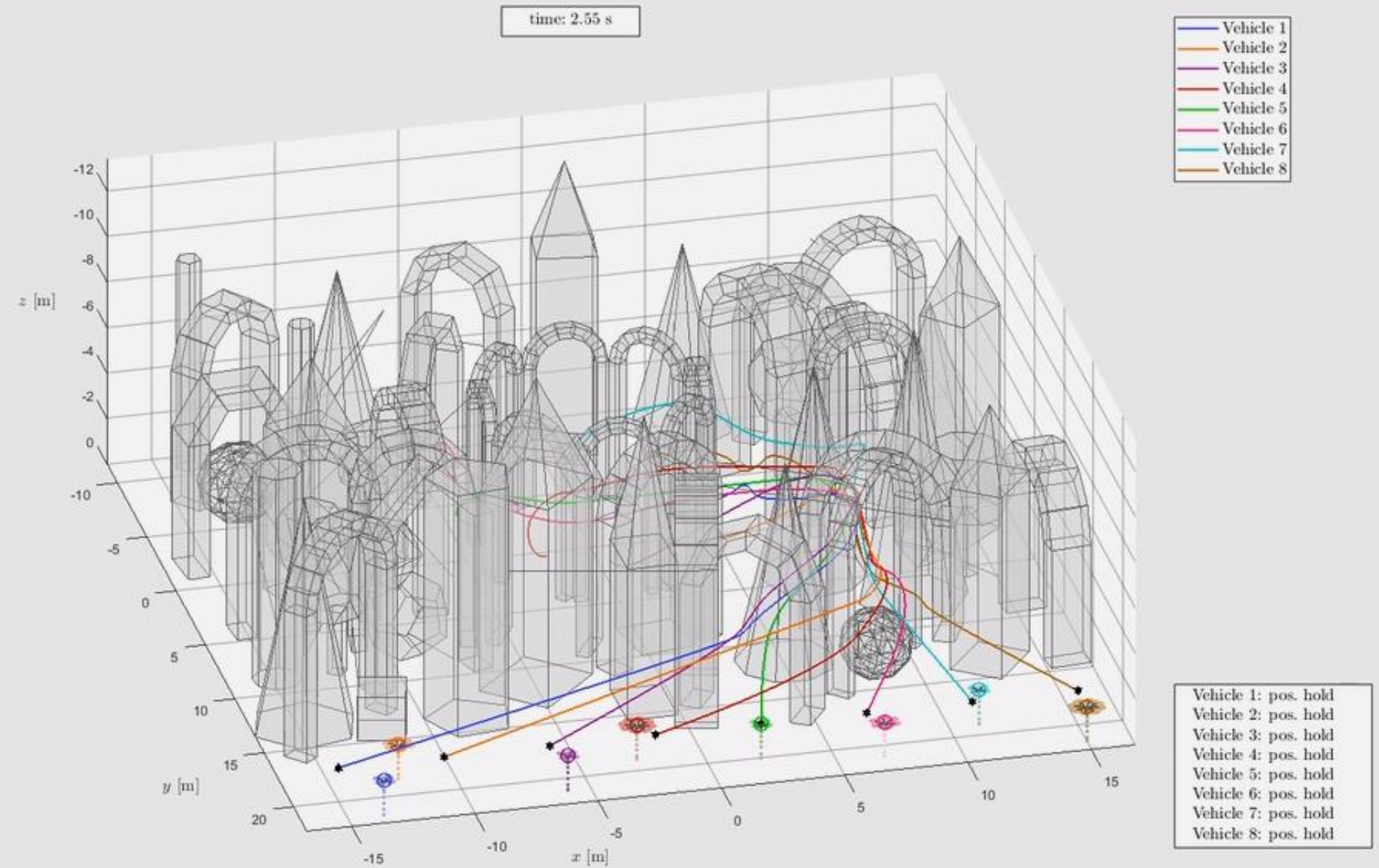
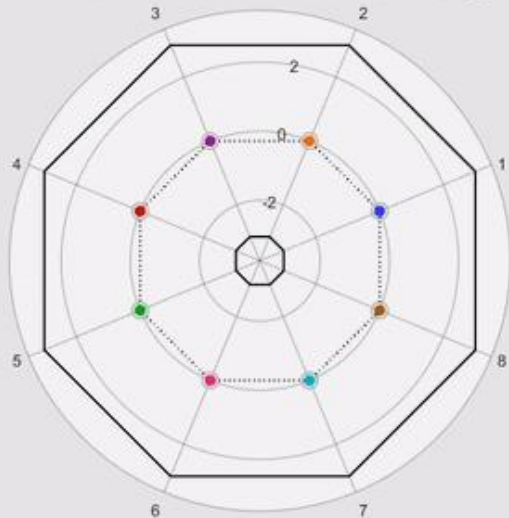
- Under reasonable assumptions (uniformly continuous trajectories, by construction; path following error, and time-critical coordination performance), safety-critical bounds are satisfied

Application: Coordinated Airspace Management

Coordination constraints: $\xi_i - \xi_j$ [s]

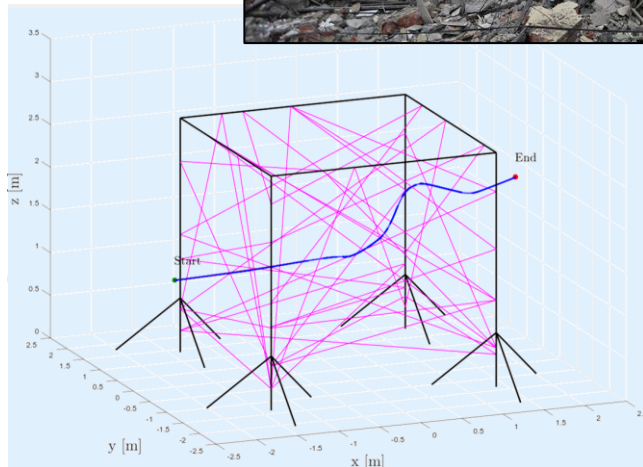


Temporal constraints: $\xi_i - \xi_R$ [s]

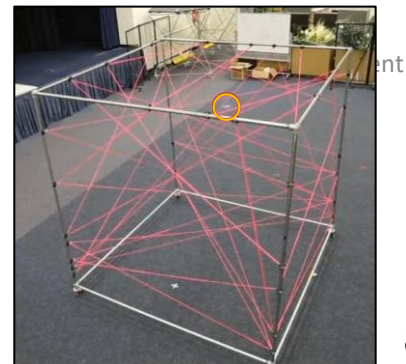


The Wire Maze

Effects of an earthquake in Ecuador



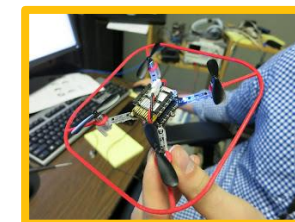
An analog for a cluttered environment in simulation...



and in reality.



Multi-agent Wire Maze Flight Demo



Crazyflie™ UAV

ATTRACTOR Component: Trustworthiness of ML-enabled Safety-critical Systems

Alwyn Goodloe

- **Findings:**

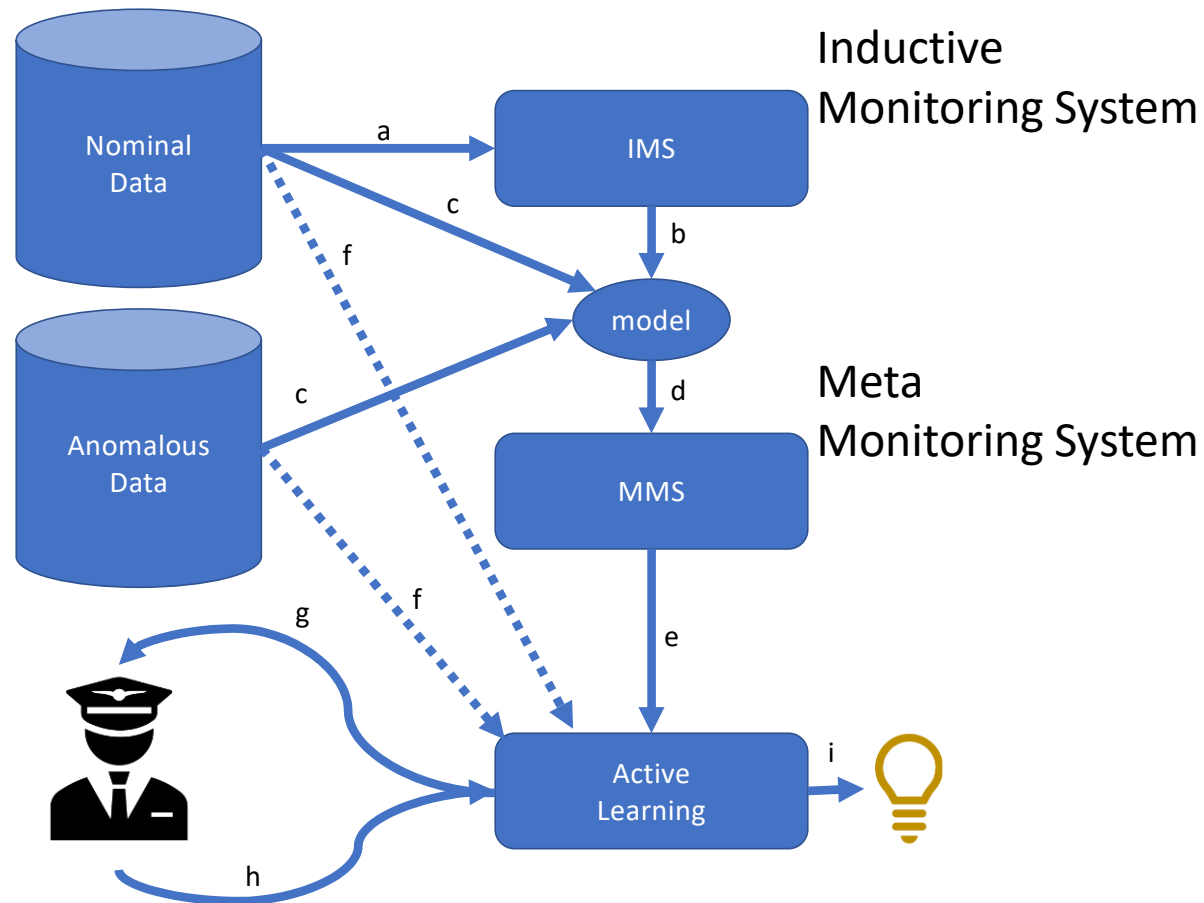
- At current SOA, it may not be possible to write specifications of the function correctness for SOA ML systems. This makes them not amenable to traditional software verification techniques.
- Basic research remains to be done on how to verify functional correctness of ML enabled systems.
- Researchers are investigating verifying correctness properties of neural networks.
- For the foreseeable future, runtime verification will have to be used extensively. However, runtime verification faces the same challenge as formal verification approaches: the need to specify the property being verified.
- Perception poses a special problem because properties amenable to formalization are highly specialized and have yet to be invented for perception.

Incorporating Human Knowledge in Autonomous Systems via ML

Nikunj Oza, Kevin Bradner, David Iverson, Adwait Sahasrabhojane, Shawn Wolfe

- Idea/Concept: Facilitate trust in autonomous systems by using domain expert knowledge to identify anomalies and their precursors. explain anomalies during operations.

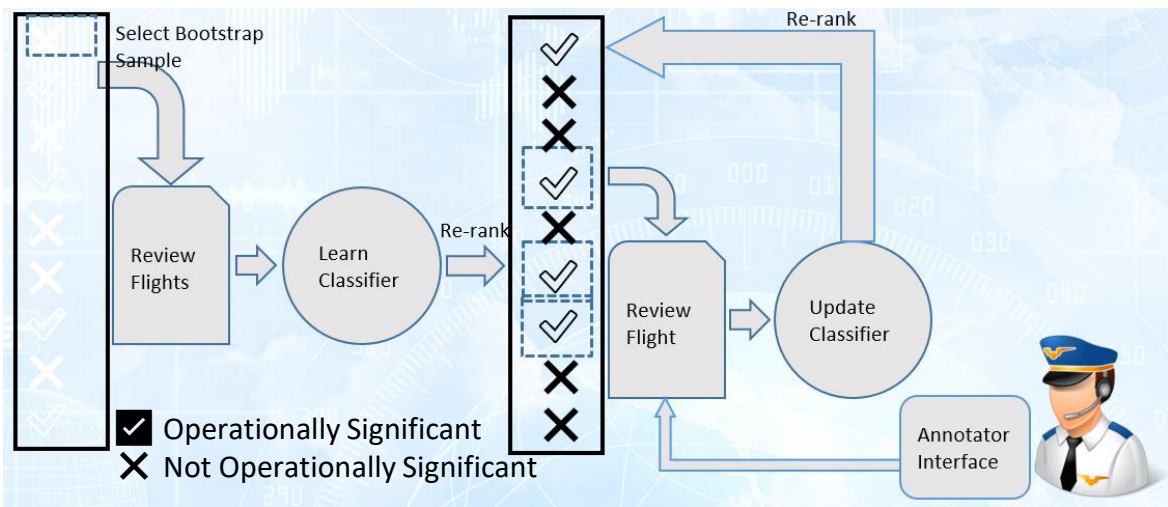
Anomaly Detection Pipeline



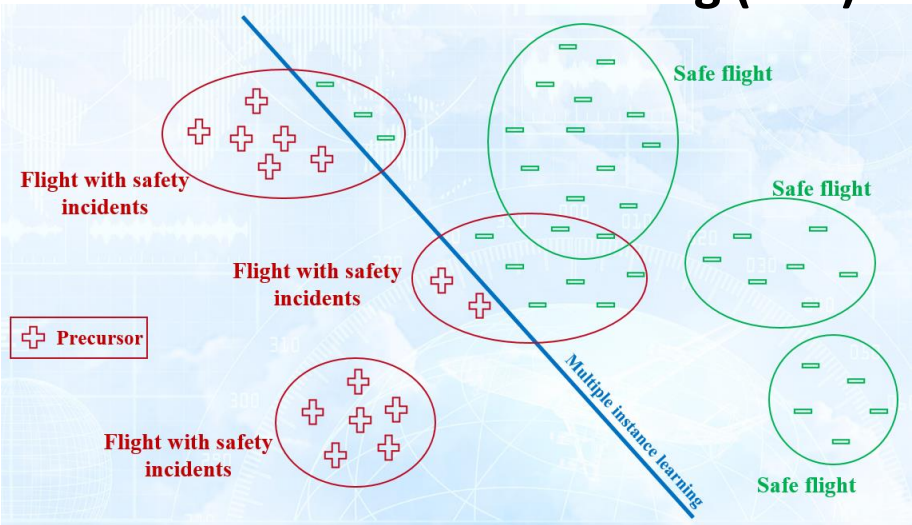
- IMS deviation scores need interpretation
- MMS post-processes IMS scores
- MMS score evaluates probability that each observation was generated from an off-nominal system

Anomaly Detection, cont.

Active Learning



Multi-instance Learning (MIL)

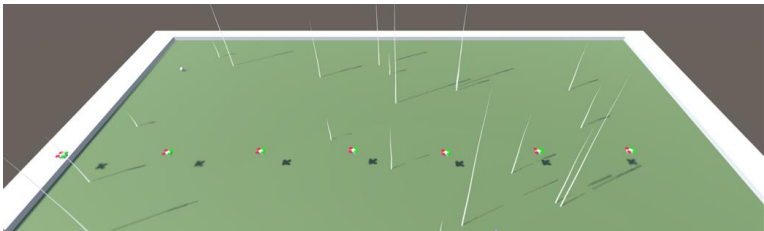


Testing

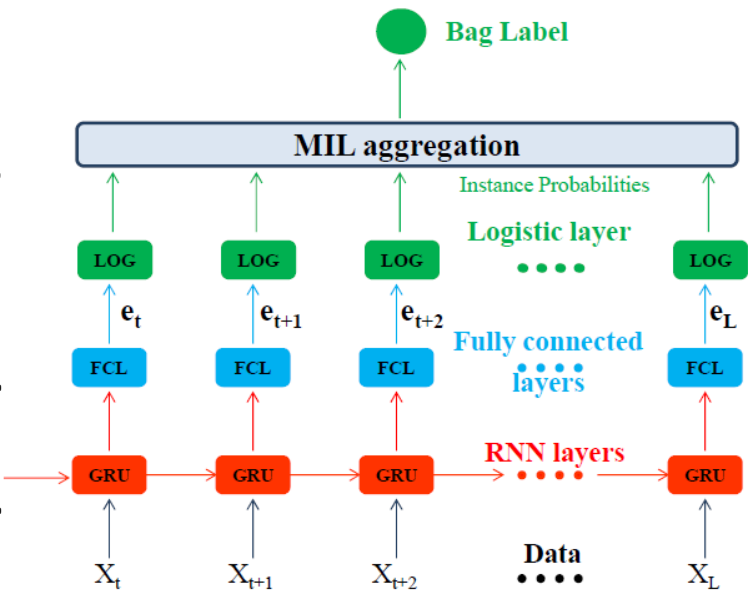
- 100 simulations of each failure type (rotor failure, tree collisions)
- Varied search area, number of drones, operating characteristics
- Train and test within scenarios, across scenarios

Findings:

- Successful detection of operationally significant vs. non-operationally significant anomalies
- Successful identification of precursors to anomalies



Deep Temporal MIL (DT-MIL)

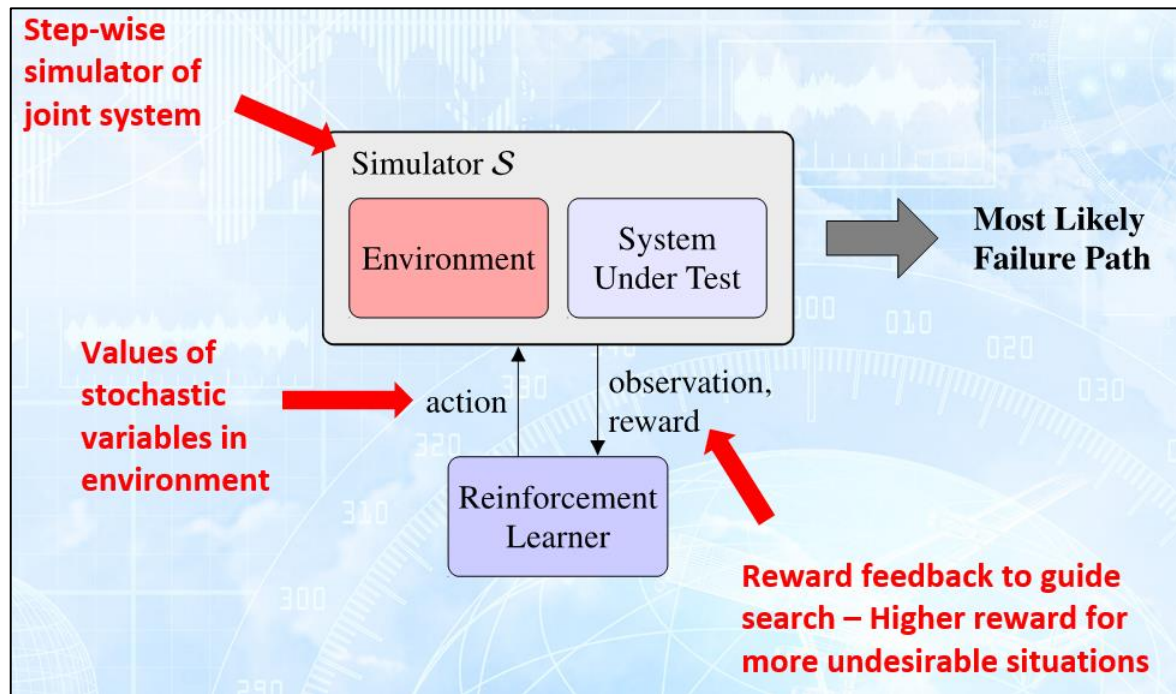


Adaptive Stress Testing (AST) of Trajectory Planning Systems

Ritchie Lee, Javier Puig-Navarro, Adrian Agogino,

Dimitra Giannakopoulou, Ole Mengshoel, Mykel Kochenderfer, Danette Allen

- Idea/Concept: Validation of Trajectory Planners (TP) using Adaptive Stress Testing (AST) to increase confidence in TPs

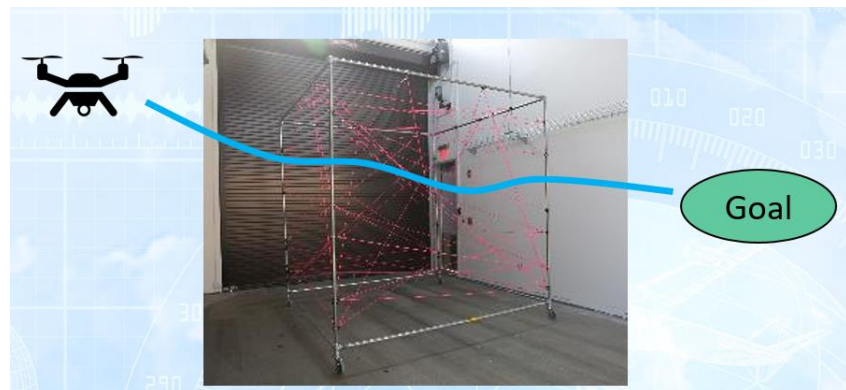


SOA (Monte-Carlo) does not scale

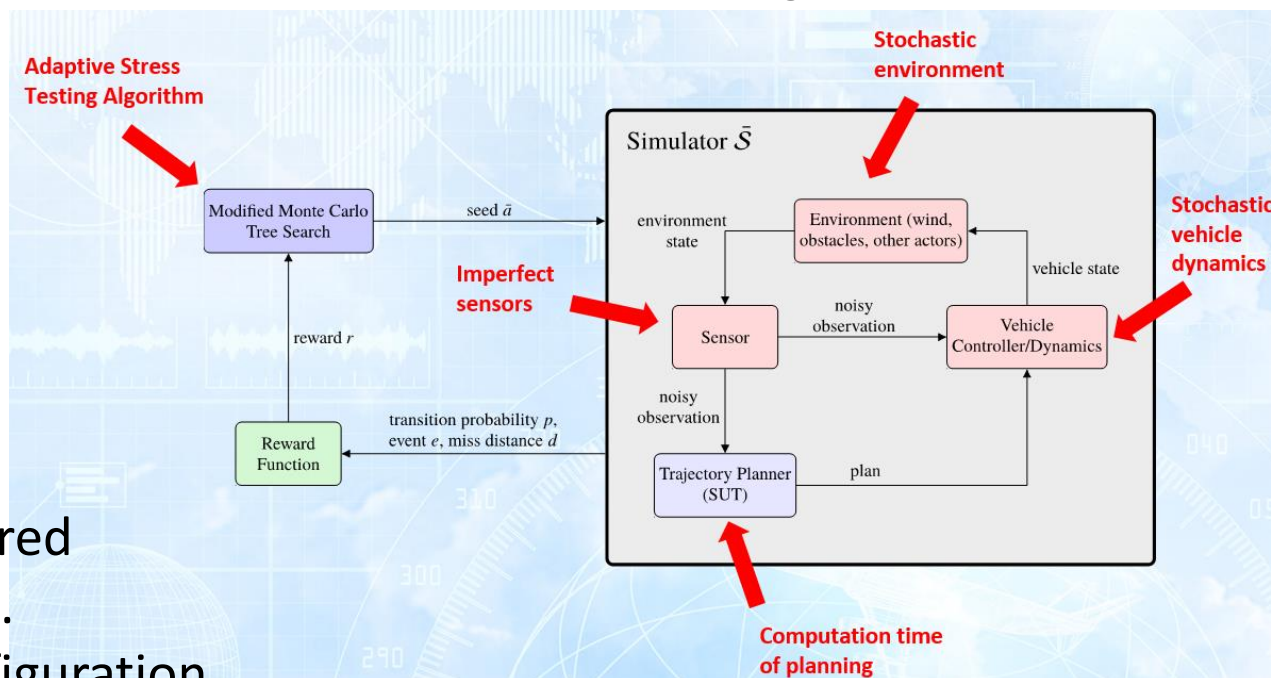
AST

- Formulates finding the **most likely failure path** as a **sequential decision-making problem**
- Apply reinforcement learning algorithms for finding failures
- Successful applications in Aircraft collision avoidance systems (ACAS X) and self-driving cars

AST: Application to Wire Maze Scenario



Testing Architecture



Experiments:

- Stressed for collision with obstacles; considered planning failures, implementation issues, etc.
- Testing dimensions: Sensor noise, initial configuration, replanning condition, computation time, dynamics noise

Results:

- 1600 total scenarios searched
- 386 successfully reached goal
- 19 collisions
- 1195 terminated in planning failure

Findings:

- AST can be extended and scaled to trajectory planning
- AST can run unanticipated scenarios and find a variety of errors
- Helped understand the residual risk, failure modes of the system
- Increased understanding of system's trustworthiness

A Critical Aspect of Uncertainty (NMA)



- Agents must arrive at satisfactory decisions in time.
- Reducing solvability uncertainty is always done via bounding problem complexity; e.g., current airspace.
- Q: in a dynamic human-machine team, which agent should have decision authority at time t ?
- Idea
 - Represent problem complexity in a computable form.
 - Build models of problem solution behaviors as a function of problem complexity.
 - Detect approaches to unacceptable complexity.
 - Reconfigure system and/or decision-making.

Measure of Complexity

- Tractability of the problem solution on a time budget, as a function of external and internal parameters.
- The quality of solutions of the decision-making problem:
 - Constraint satisfaction
 - Distance from optimality
 - Robustness (maximum uncertainty under which performance is guaranteed)
 - Serendipity (minimum uncertainty that would allow for great improvements in performance)
- Other, context-dependent and formulation-dependent functions.
- **Results to date:** Simulated decision-makers show clear trends in the ability to solve decision-making problem on a budget, with varying degrees of “self-awareness”.

Broader Issues: “AI” and Cognitive Problem

- Well-known problem in human-machine interactions: unjustified overreliance or under-reliance on computational tools
- Recent findings
 - The Flynn effect is in a reverse trend
 - In education: significant negative impact of “AI” on decision-making and safety
- Not so recent but increasing effect: decline of design expertise in industry
 - Push-button use of AI and simulations does not develop knowledge or discernment necessary for sound decision-making
- **Urgent gap: How to use AI without adverse effects on human cognition?**

Concluding Remarks

- Key question: In a safety-critical system, what constitutes sufficient evidence that an ML agent or system component meets its requirements?
- Traditional solutions still apply
 - Risk-based design
 - In the case of irreducible and unknown uncertainties, design external safety measures



Khrystinasnell, CC0, via Wikimedia Commons

Public domain https://upload.wikimedia.org/wikipedia/commons/3/3e/Elevator_Doors.jpg



Unknown author - Copie de gravure ancienne

Elisha Otis demo of his free-fall prevention mechanism, Crystal Palace, 1854.

Public domain https://en.wikipedia.org/wiki/Elisha_Otis#/media/File:Elisha_OTIS_1854.jpg