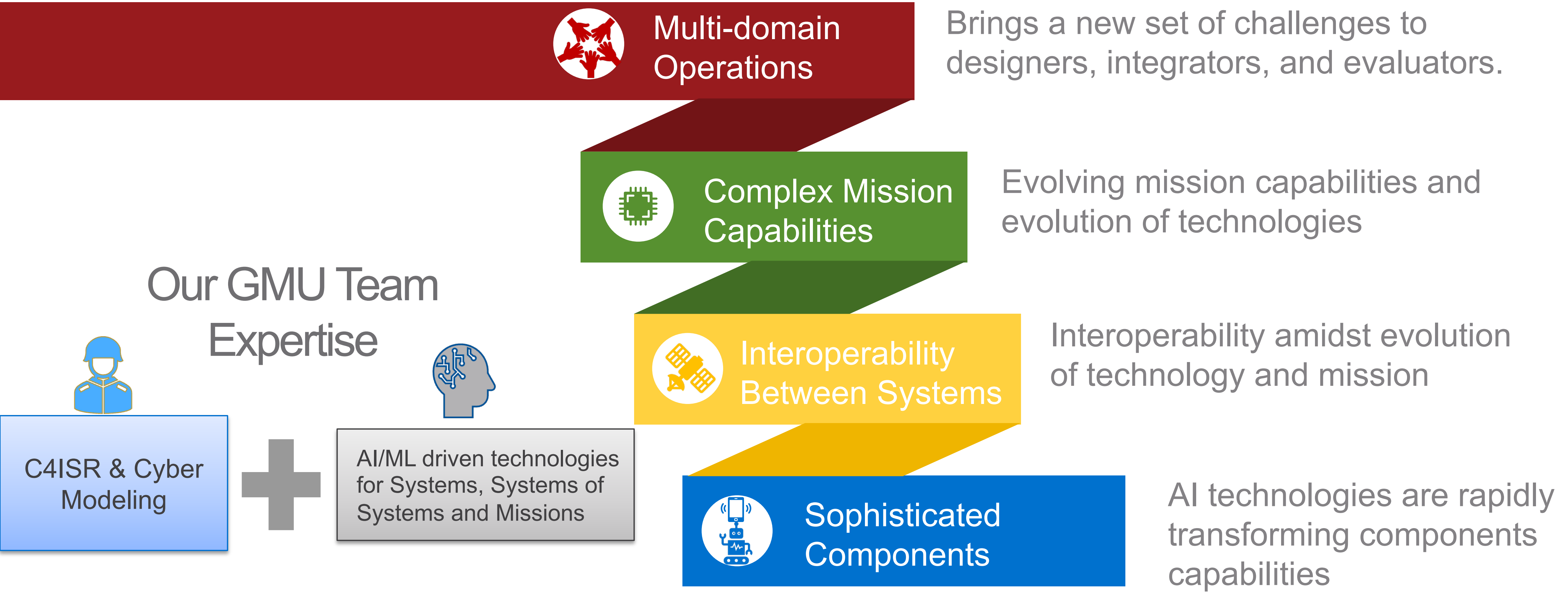
32nd Annual INCOSE international symposium
hybrid event
Detroit, MI, USA
June 25 - 30, 2022

Dr. Ali Raz, Dr. Shou Matsumoto, Dr. Paulo Costa

Exploring Explainable Artificial Intelligence to aid Systems Engineers in Design and Evaluation of Complex Systems

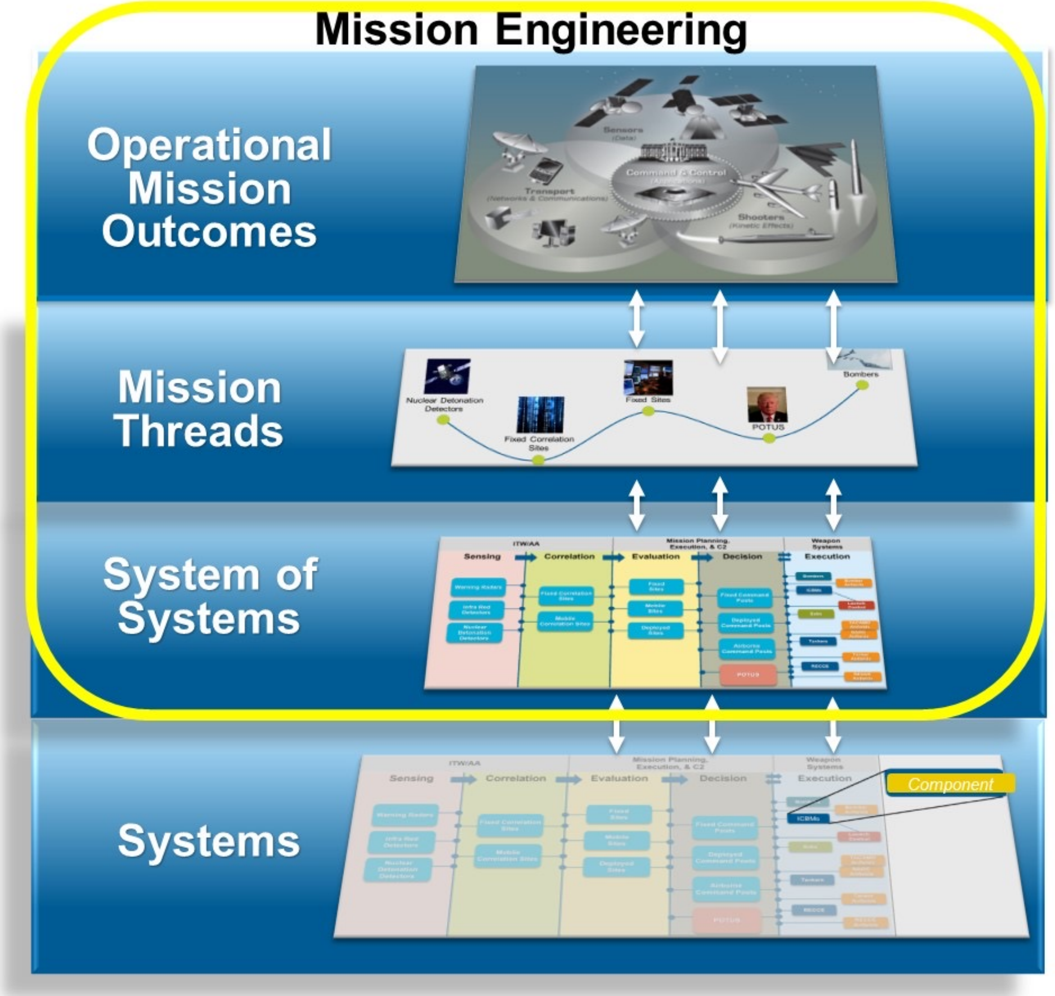
George Mason University
C4I & Cyber Center

Challenges for Complex Systems and System of Systems



AI-Enabled Capabilities in Complex Systems

Systems Engineering applies through out



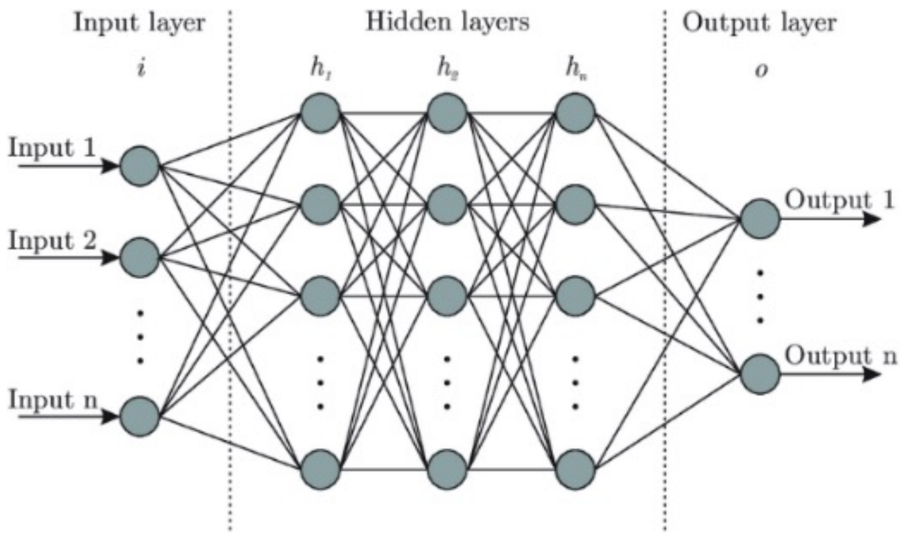
- Mission engineering (ME) is 'above the system level', addressing systems of systems (SoS) in a mission context
- AI for contextually understanding mission thread and SoS capabilities
- Recent AI Application Example: Planning battlefield operations, recommend courses of actions etc.

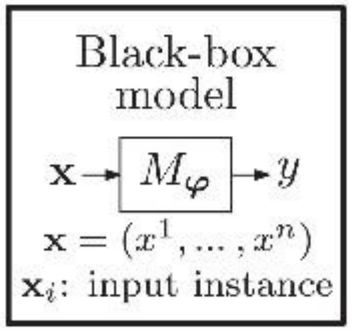
- AI for managing and integrating heterogeneous and distributed systems
- Systems with autnomous and non-autnomous systems
- Recent AI Application Example: Space Situational Awareness, Missile Defense, etc.

- AI at system, sub-system, and functional level
  (New independent systems/upgrade legacy systems)
- Recent AI Application Examples: Control, Guidance, and Navigation of Robots, Missiles, Autnomous Cars, etc.

What is the Challenge with AI-Enabled Capabilities?

- Deep Neural Networks (DNNs) remain the predominant implementation of AI and Machine Learning
- AI Systems with DNNs can **perceive, learn, decide, and act***





1. Why this action?
2. Why not another action?
3. When do I succeed/fail?
4. When can I trust the results?
5. How can I fix an error?

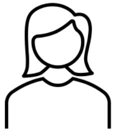
- Current AI models excel in performance, but extremely hard to interpret
  - There is a trade-off between the performance and interpretability of AI models
  - Decision making constructs are opaque which makes trust, validation, and certification difficult

- EXplainable AI (XAI) is developing techniques to address this problem

Why Explainable AI?

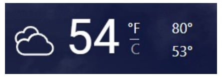
Let's try a thought experiment

Q: What will be the weather tomorrow?

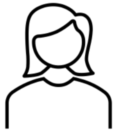


It will be chilly and cloudy.
Remember to put on a warm jacket

alexa

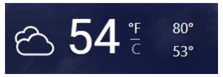


Q: How do you know what will be the weather tomorrow?



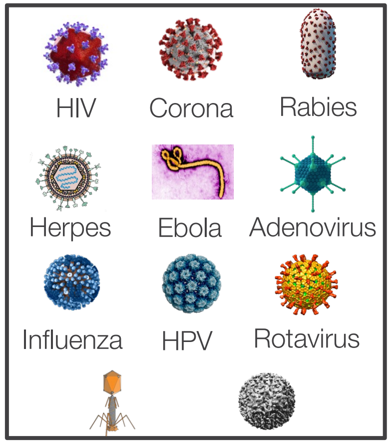
I heard it on the radio
I looked up on my phone
Weather radar showed a cold front
I love looking at NOAA models, you wanna know the barometric pressure!

alexa



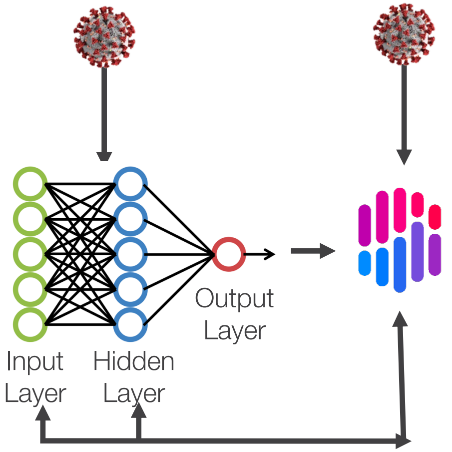
Concept of Explainable AI

ML with XAI



Training Data

Learning Process



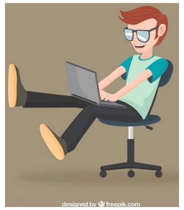
Learned Function

Explainable Model

This is Corona virus because:

1. R_0 between 2 and 3
2. Incubation 4 to 14 days
3. Droplets in the air for long
4. A pandemic

Explanation Interface



End User

1. I understand why/why not
2. I know when I succeed/fail
3. I know when I can trust the results
4. I know why there is an error

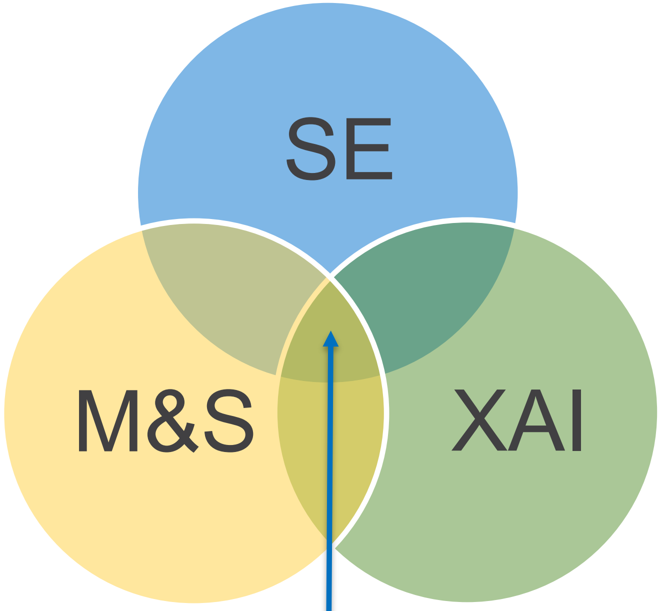
Our Approach to Complex Systems Design and AI Challenges

Systems Engineering

Complex systems are built on SE foundations

Modeling and Simulation

Use models as a basis for simulations to develop data for analysis, education, decision making, *etc*.

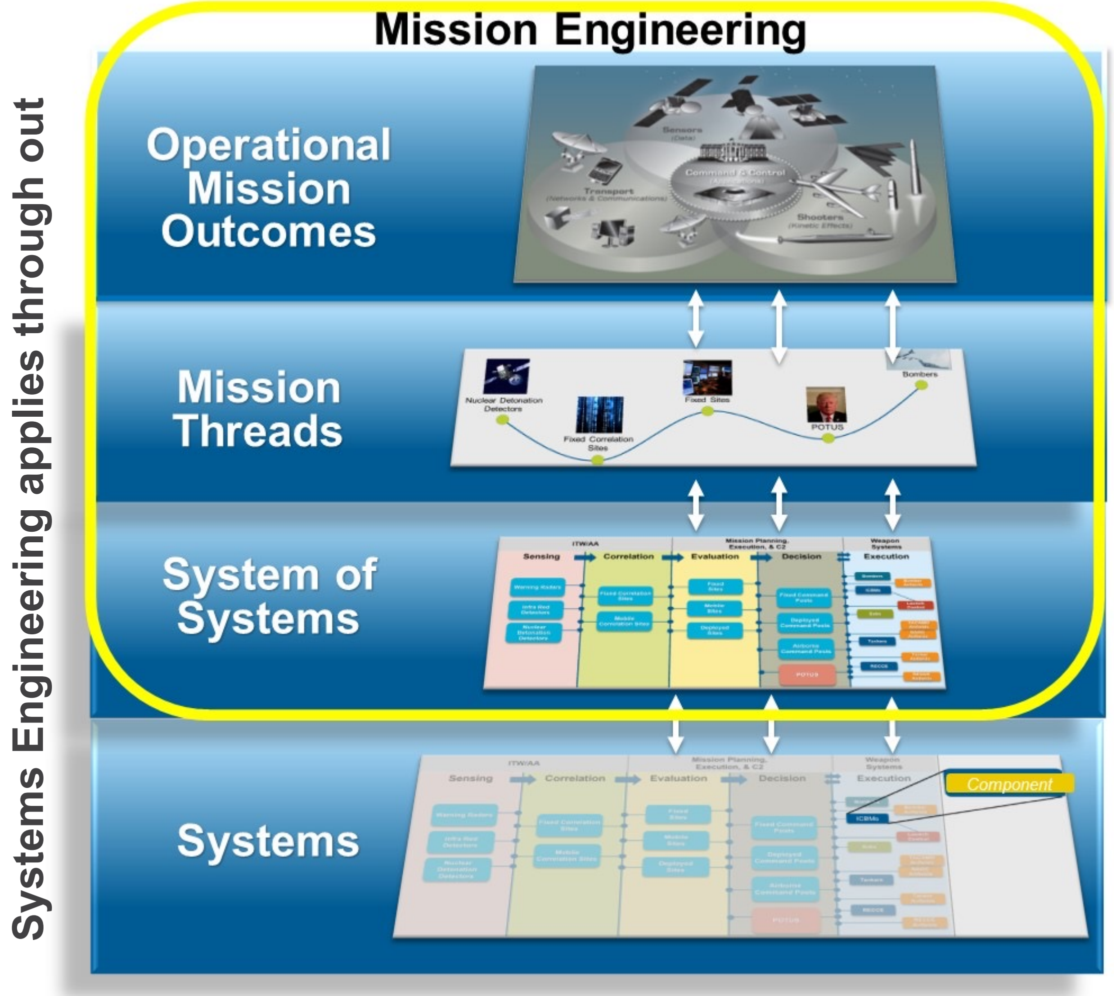


Explainable Artificial Intelligence

AI in which the solution (including results, rationale, justification, assumptions, *etc.*) can be interpreted and understood by humans.

Provide a window of opportunity to understand, analyze and validate the assumptions, theories, operations, and decision-making constructs of modern complex systems.

Applying XAI to aid with Systems Engineering of Complex Systems



- Problem: Creating balance and unbalance in complex operational environments
- AI Methods: Convolutional Neural Networks
- XAI Methods: Gaussian Processes with uncertanity

- Problem: Machine reasoning for integration of system of systems
- AI Methods: Hierarchical Task Networks
- XAI Methods: Probabilitic Ontology with Multi-Entity Bayesian Networks

- Problem: Guidance of high-speed aerospace systems
- AI Methods: Reinforcement Learning
- XAI Methods: Shapley Additive Explanations

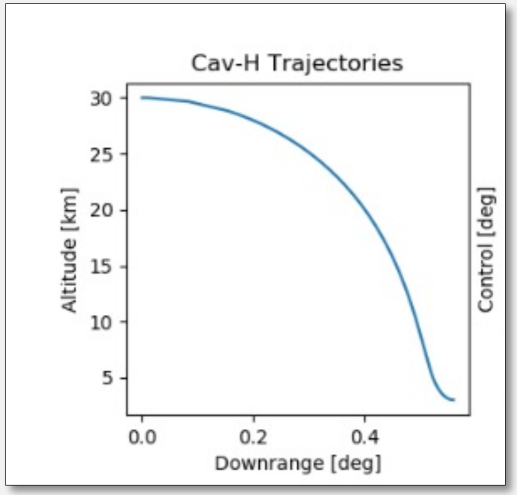
System Level AI: Guidance and Navigation of High-speed Aerospace Vehicles

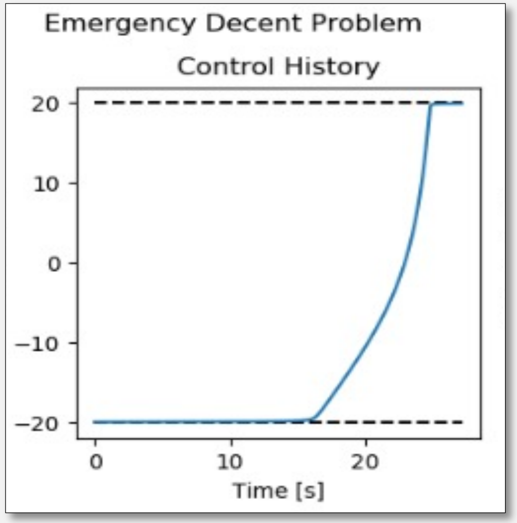
- **Problem Formulation:**
  - Provide guidance commands to high-speed aerospace vehicles

- **Why AI is Needed:**
  - Real-time trajectory generation is computationally prohibitive for high-speed mission
  - Human intervention and guidance is not feasible beyond supersonic speeds

- **What AI Techniques can be used:**
  - Reinforcement Learning (RL) enables training of an artificial intelligence (AI) agent to operate in dynamic uncertain environments
  - Impressive performance outcomes to learn nearly-optimal solutions in a variety of application domains

Sample Problem





Emergency descent of vehicle from 30km altitude to 3km altitude in minimum time

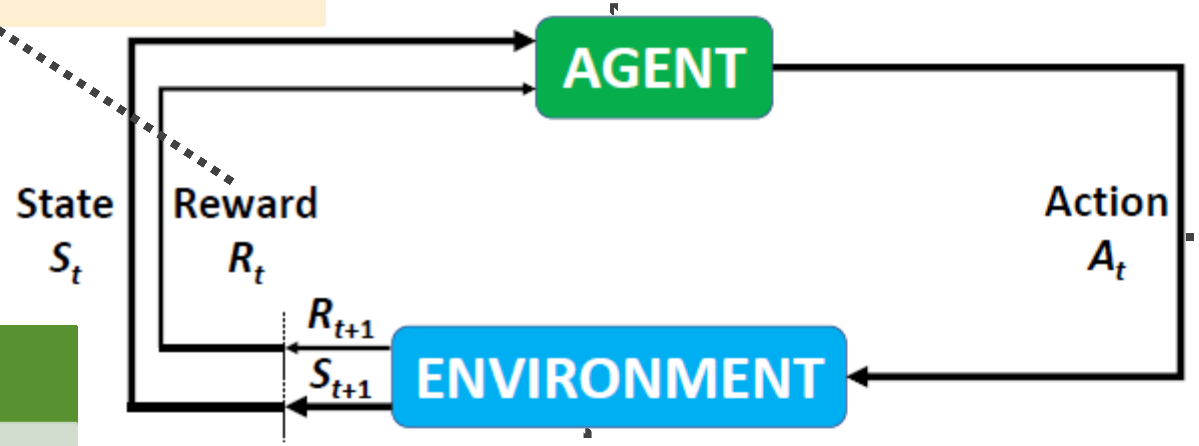
Reinforcement Learning Problem Formulation

Reward Function

- Designed to train the RL agent an emergency descent problem
- Reward structure based on distance to target and FPA

RL Agent

- RL trained from SB3 Python package
- Proximal Policy Optimization (PPO)
- RL training parameters (backup)



Action Space

- AoA command
- ±20º in variable increments of 2º

State

- Distance to target
- Altitude
- Velocity
- AoA
- FPA

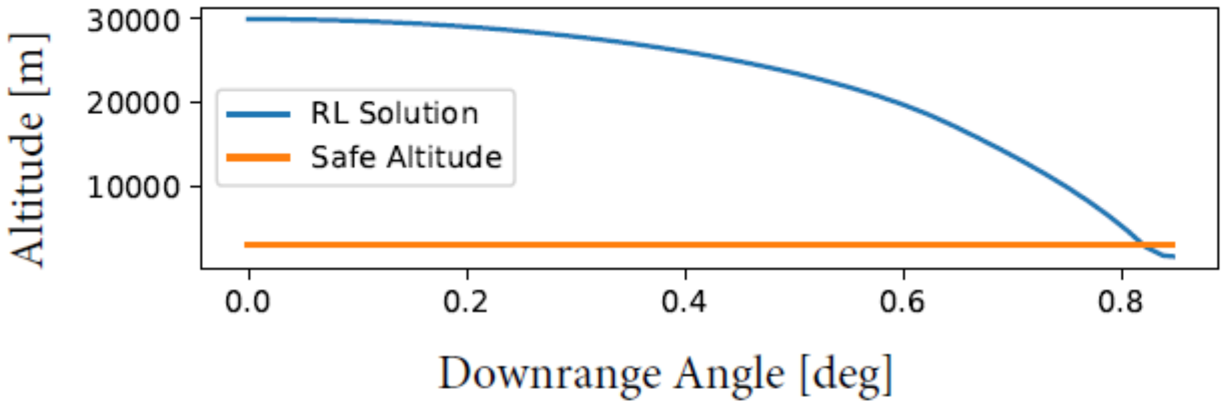
Environment

- Emergency descent problem space
- Atmosphere, simulation clock, scheduler, etc.

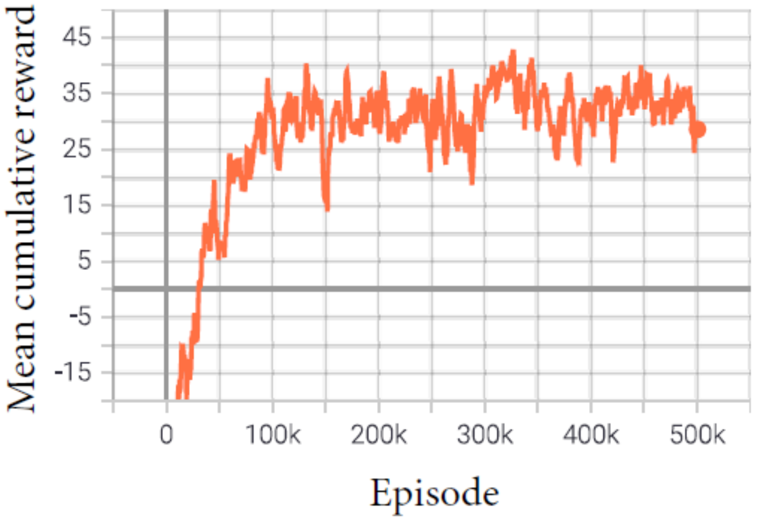
AI Result (Vehicle Descent From 30 km to 3 km)

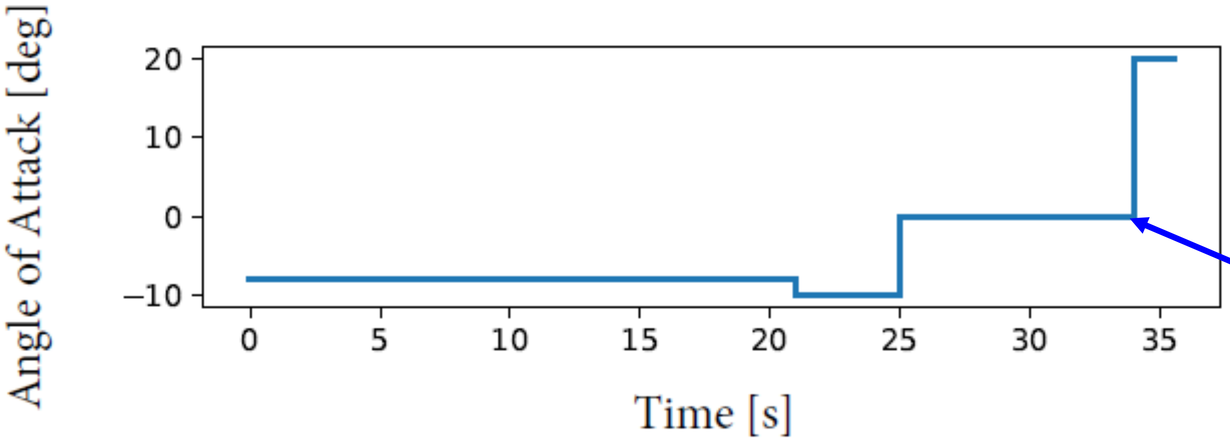
AI based on Reinforcement Learning:

- Provides AoA commands to guide the vehicle to a pre-determined safe altitude
- Included randomly sampling vehicle initial conditions
- Completed after 500k episodes



Sufficient cumulative reward of +30 to train policy





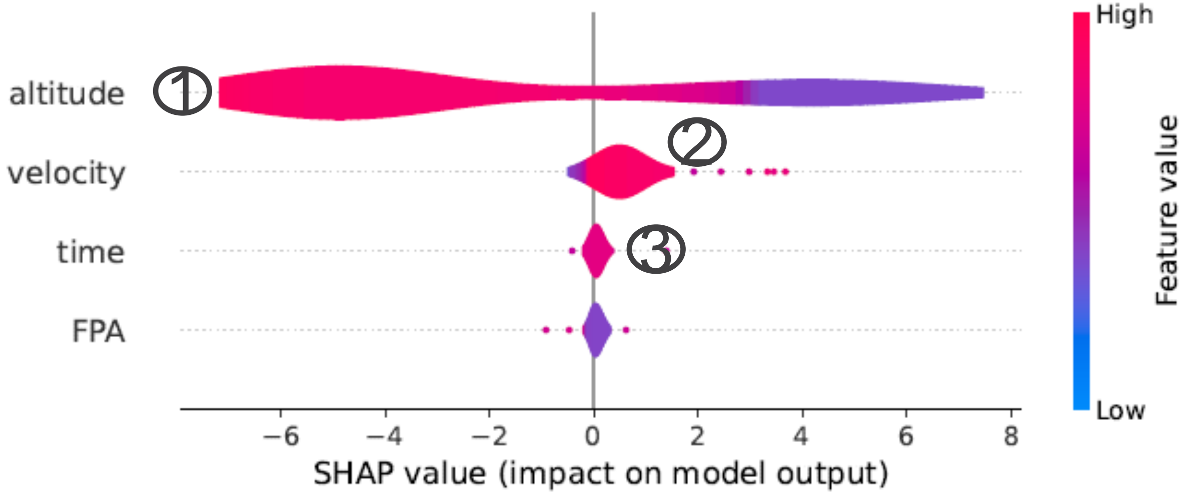
Examination Via Explainable AI (XAI) Techniques

Brief Introduction to Explainable AI

- Investigates trained Deep Neural Network (DNN) models with analytical techniques to extract decision making attributes
- SHapley Additive exPlanations (SHAP)
  - State of the art for reverse engineering the output of any predictive model
  - Yields importance of input features for a given prediction
  - Focuses on coalitions in cooperative game theory

SHAP Applied to RL Problem

- **Inputs:** Time, Altitude, Velocity, and Flight Path Angle
- **Output:** Angle of Attack (between -20˚ and 20˚)
- **Number of Trajectories:** 1000
- **Objective:** Reach a particular target in a minimum time



1. Higher altitude values oppose a change in AoA whereas lower altitudes support it.
2. Higher velocity values positively influence change in AoA
3. FPA and Time have least impact.

System of Systems Level AI: Anytime Reasoning and Analysis for Kill-Web Negotiation and Instantiation across Domains (ARAKNID)

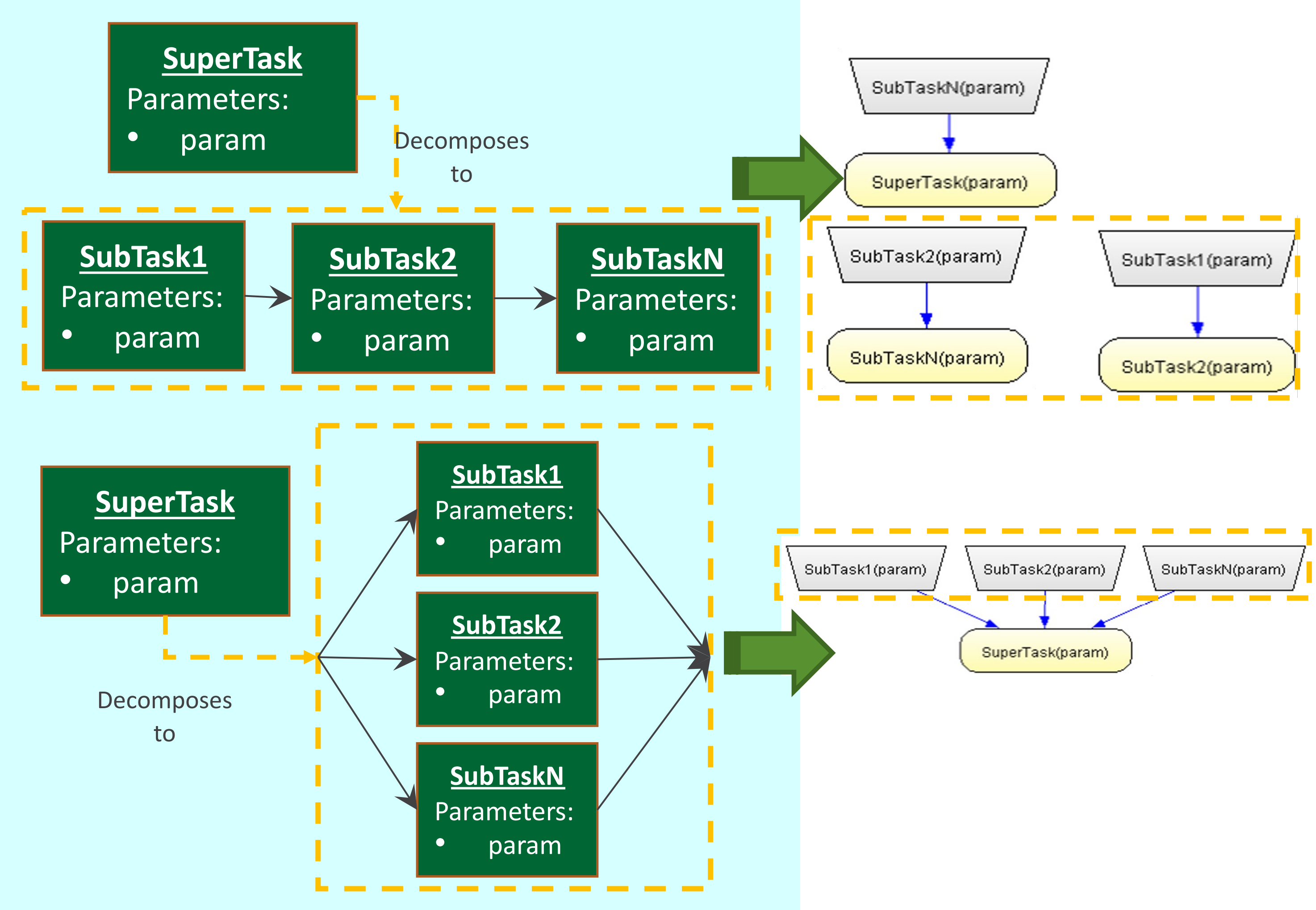
- **Problem Formulation**
  - Near real-time flexible and adaptive integration of system of systems
  - Part of the DARPA's *Adapting Cross-domain Kill-webs* (ACK) program and Mosiac Warfare Concepts
    - https://www.darpa.mil/program/adapting-cross-domain-kill-webs

- **Why AI is needed:**
  - Manage data, capabilities, and outcomes on various missions in real-time to help with decisions involving highly complex and capable complex systems

Automating machine-to-machine interactions is a significant tactical advantage when seconds and minutes matter *(Important decisions are ultimately made by a real person)*

- **What AI and XAI techniques can be used:**
  - Hierarchal Task Networks for mission planning
  - Combining Probabilistic Ontologies with Multi-entity Bayesian Networks to provide explainable AI insights into the design of adaptable system of systems architectures

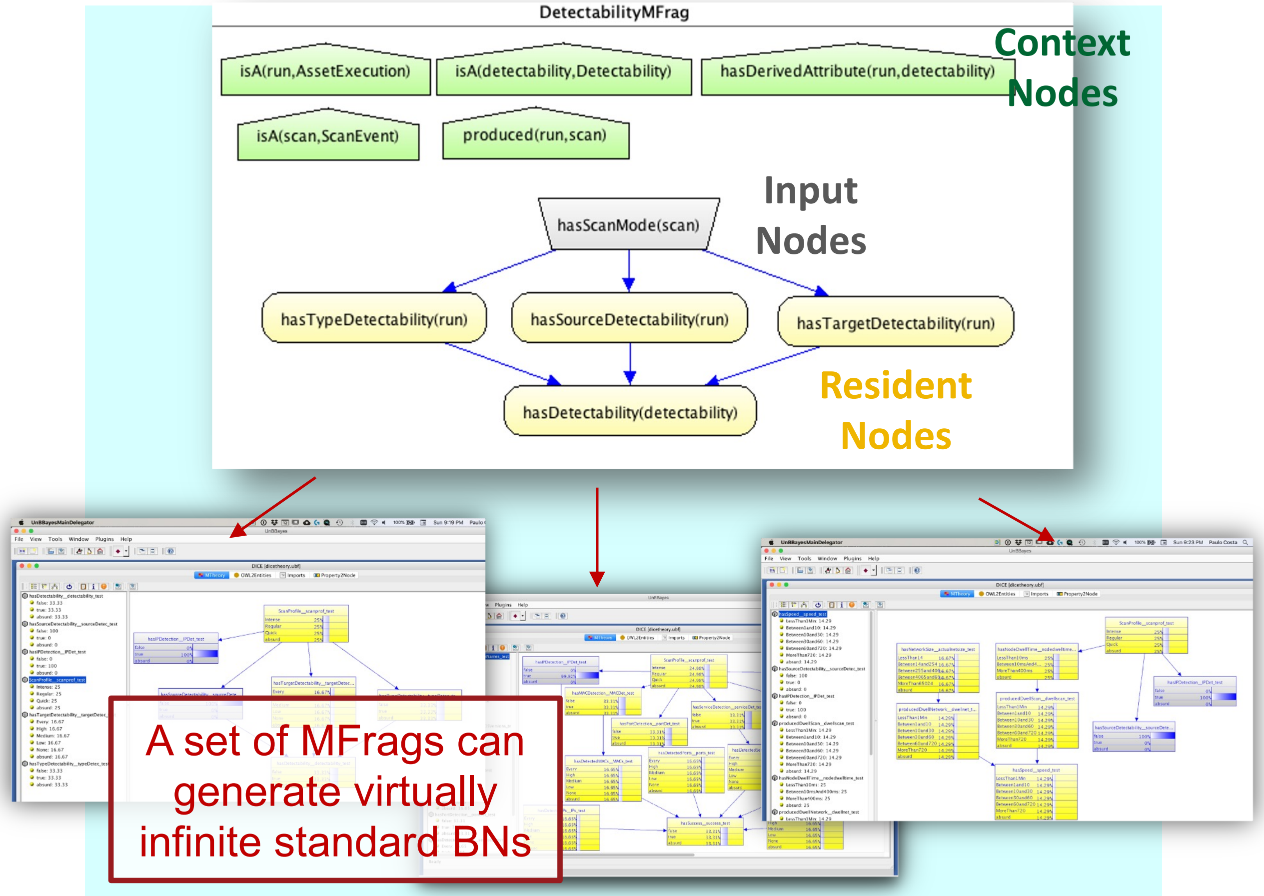
AI Technology: Hierarchical Task Networks (HTN)



Hierarchical Task Network (HTN):

- Applied in automated planning
- Represents task (action) flows
  - Dependency among actions form a "task → subtask" hierarchy.
    - Produces executable sequence by decomposing tasks to subtasks
  - Preconditions and post-conditions (*a.k.a.*, constraints) can be specified with predicates & logic expressions

Multi-Entity Bayesian Networks (MEBN) and *Probabilistic Web Ontology Language* (PR-OWL)



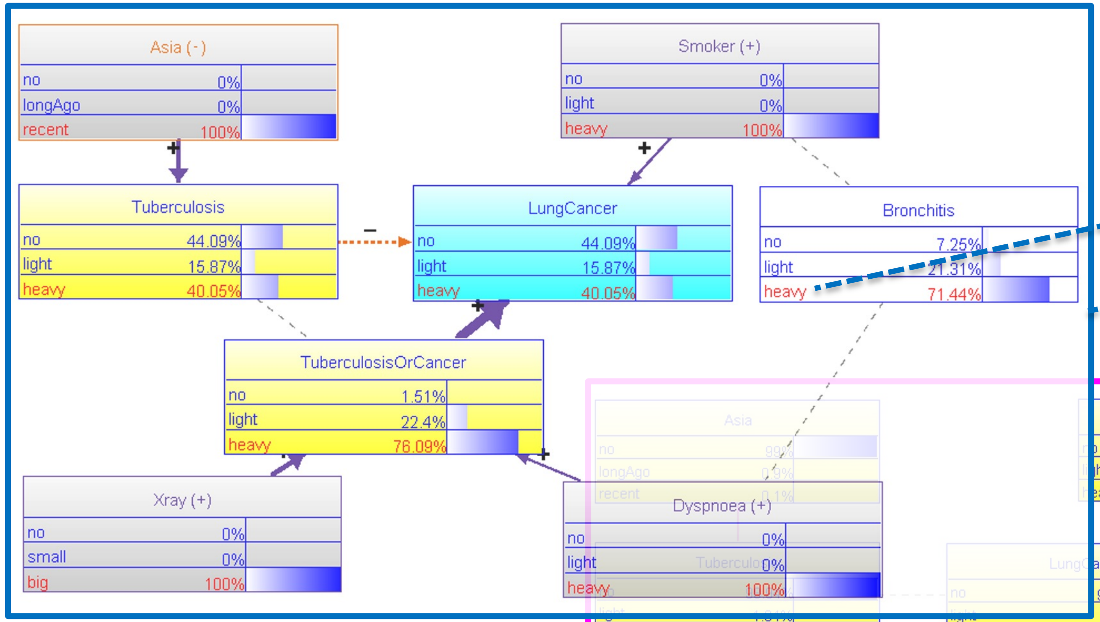
UnBBayes: https://sourceforge.net/projects/unbbayes/

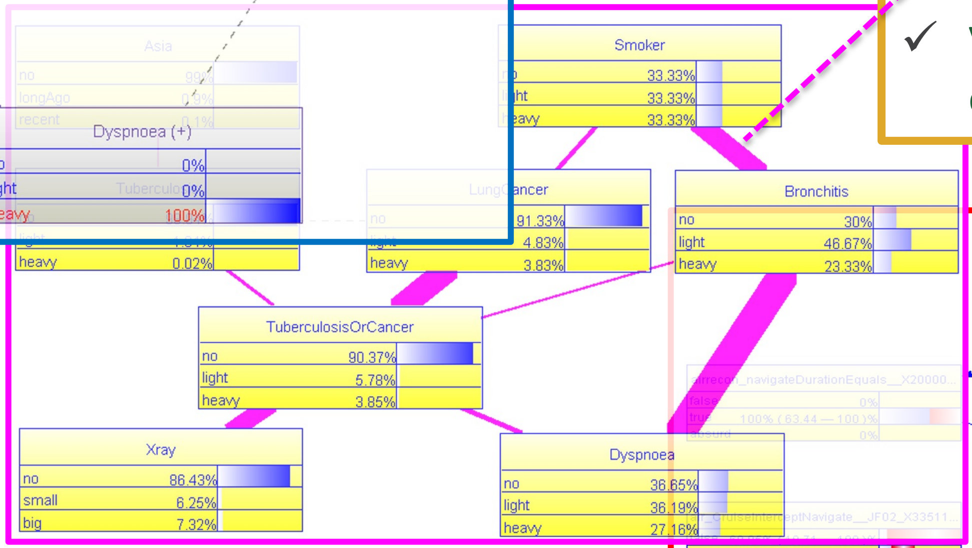
- *Bayesian Network (BN)*: directed acyclic graph & factored representation of a joint probability distribution.
- *MEBN*: composable templates for BNs based on First-Order Logic (FOL)
  - Knowledge is encoded as set of MEBN Fragments (MFrags) that can be instantiated & combined to form standard BNs for a specific situation
  - Generalize knowledge and apply to never-before-seen situations
- *PR-OWL*: language for modeling uncertainty in OWL based on MEBN semantics
  - OWL: W3C recommendation for ontologies
  - Ontology: formal specification of domain knowledge

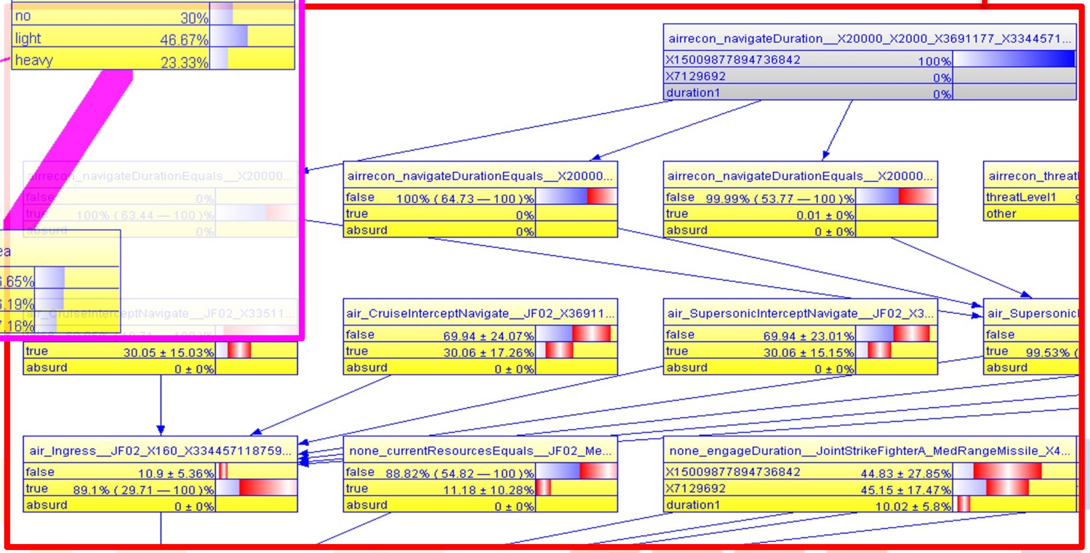
Visual Qualitative Dynamic Explanation of Bayesian Networks

"*Dynamic*" explanations are generally about <u>behaviors</u>, <u>changes</u>, and <u>reasoning processes</u>

- ✓ Abduction: Maximum a Posteriori Estimation
  - ❑ Find most probable states given evidence
- ✓ Qualitative propagation of **positive**/**negative** effects from evidence to a variable of interest.
- ✓ Visualization of strength of influence (sensitivity).
- ✓ Visualization of **intervals of probability** to show epistemic uncertainty







Mission Engineering AI: Exploiting Complex Environments to Create Advantage

- **Problem Formulation**
  - Understanding how to win in Mosaic warfare via Real Time Strategy (RTS) Games + AI
  - How do new capabilities, rules, and modifications affect winning outcomes?

- **Why AI is needed:**
  - How to account, model, and plan for intelligent opponents?
  - Discover winning strategies amidst adaptive and intelligent agents

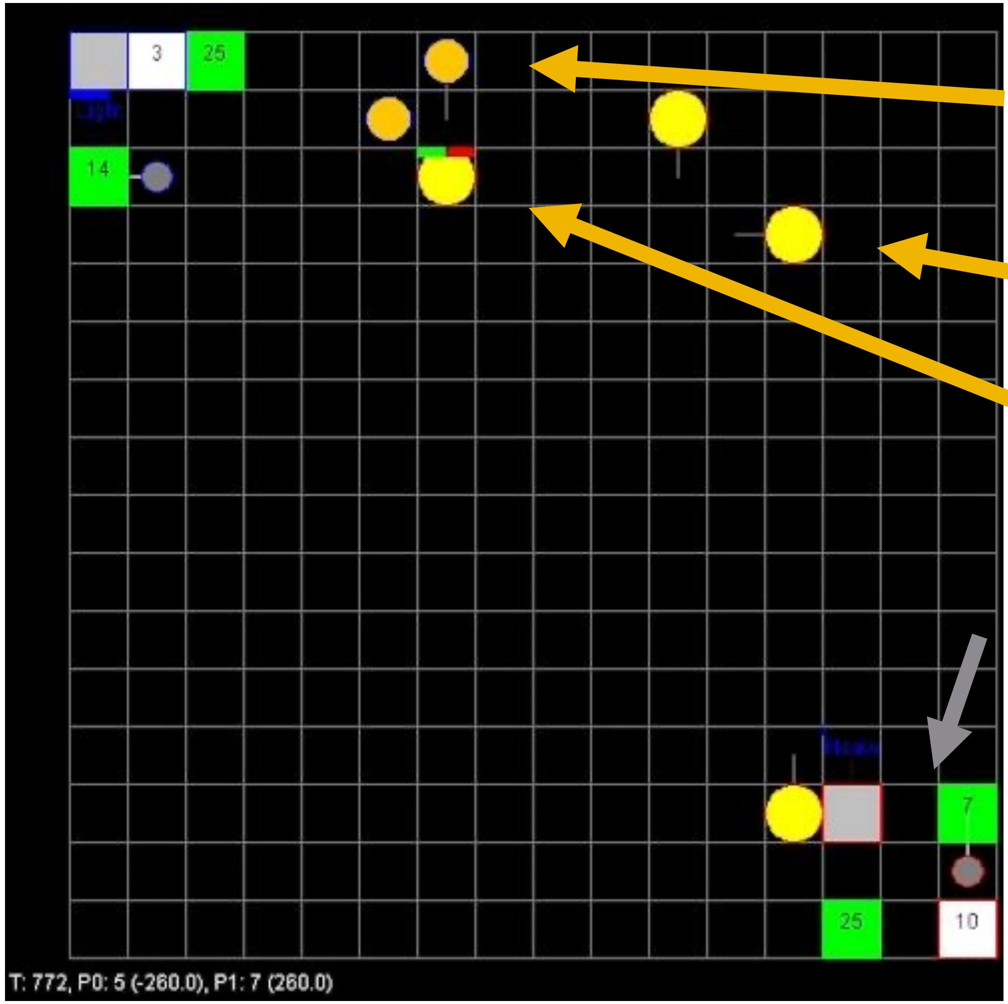
- **What AI and XAI Techniques can be used**
  - Convolution Neural Networks (CNNs) to understand enviorment
  - SHAP to investigate CNNs
  - Gaussian Processes to account for uncertanity

Based on DARPA's Gamebreaker Project in collaboration with Purdue University

microRTS: Game Specification

Distribution A: Approved for Public Release, Distribution Unlimited

Game Map



Light Unit
(Fast, Low Damage)

Heavy Unit
(Slow, High Damage)

Damaged Heavy Unit

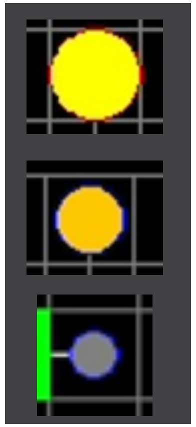
Barracks (Builds military units with resources)

Resources

Worker (Gathers resources)

Base (Builds workers with resources)

Units



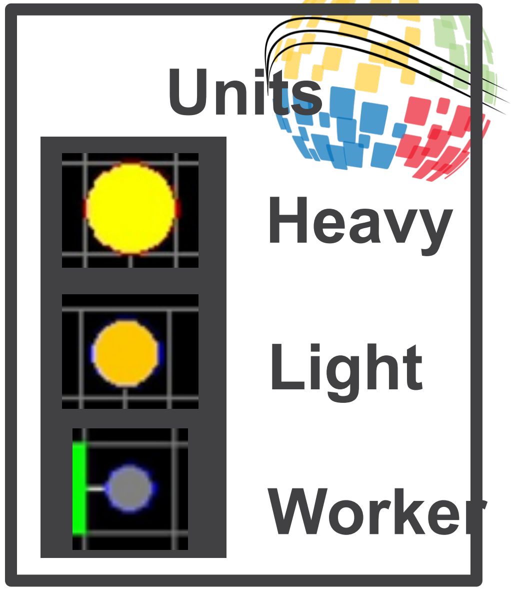
Heavy

Light

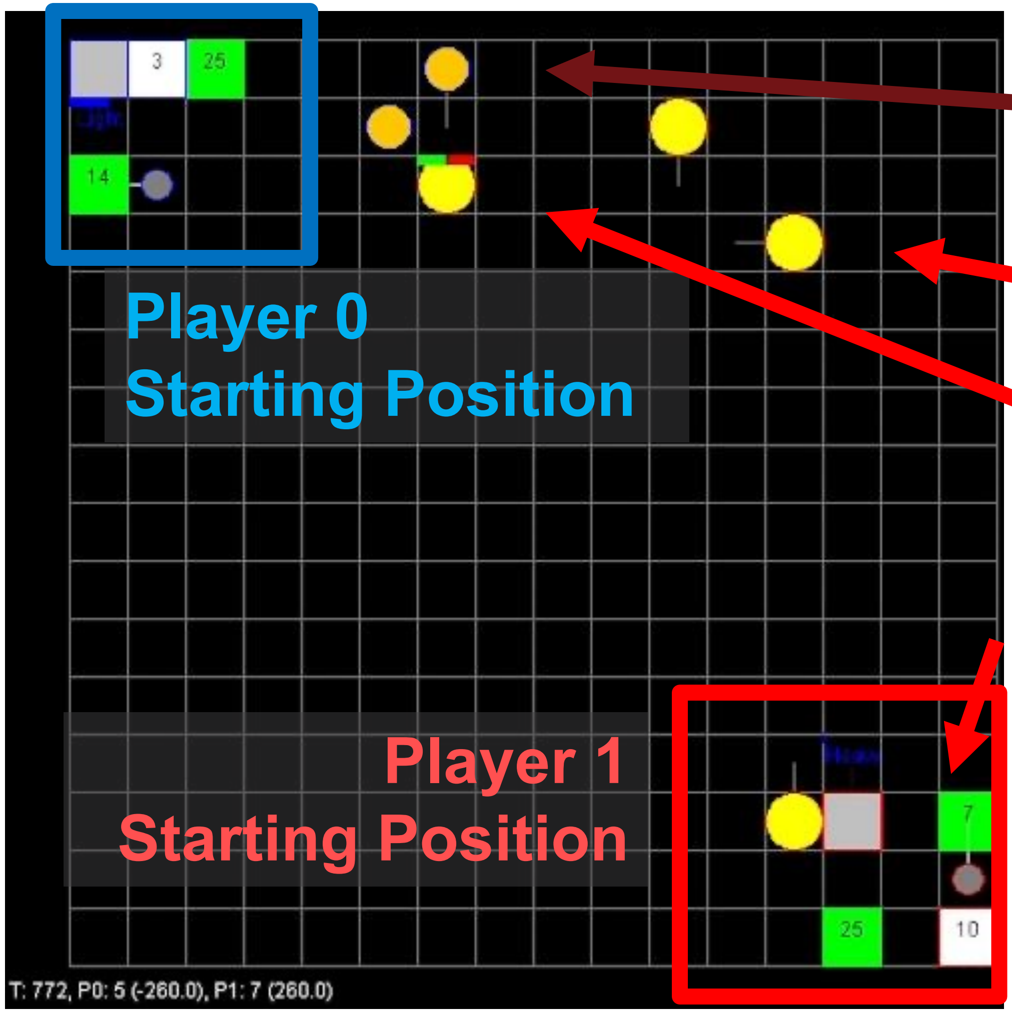
Worker

microRTS: Game Specification (Cont...)

Distribution A: Approved for Public Release, Distribution Unlimited



Game Map



Light Unit
(Fast, Low Damage)

Heavy Unit
(Slow, High Damage)

Damaged Heavy Unit

Barracks (Builds military units with resources)

Resources

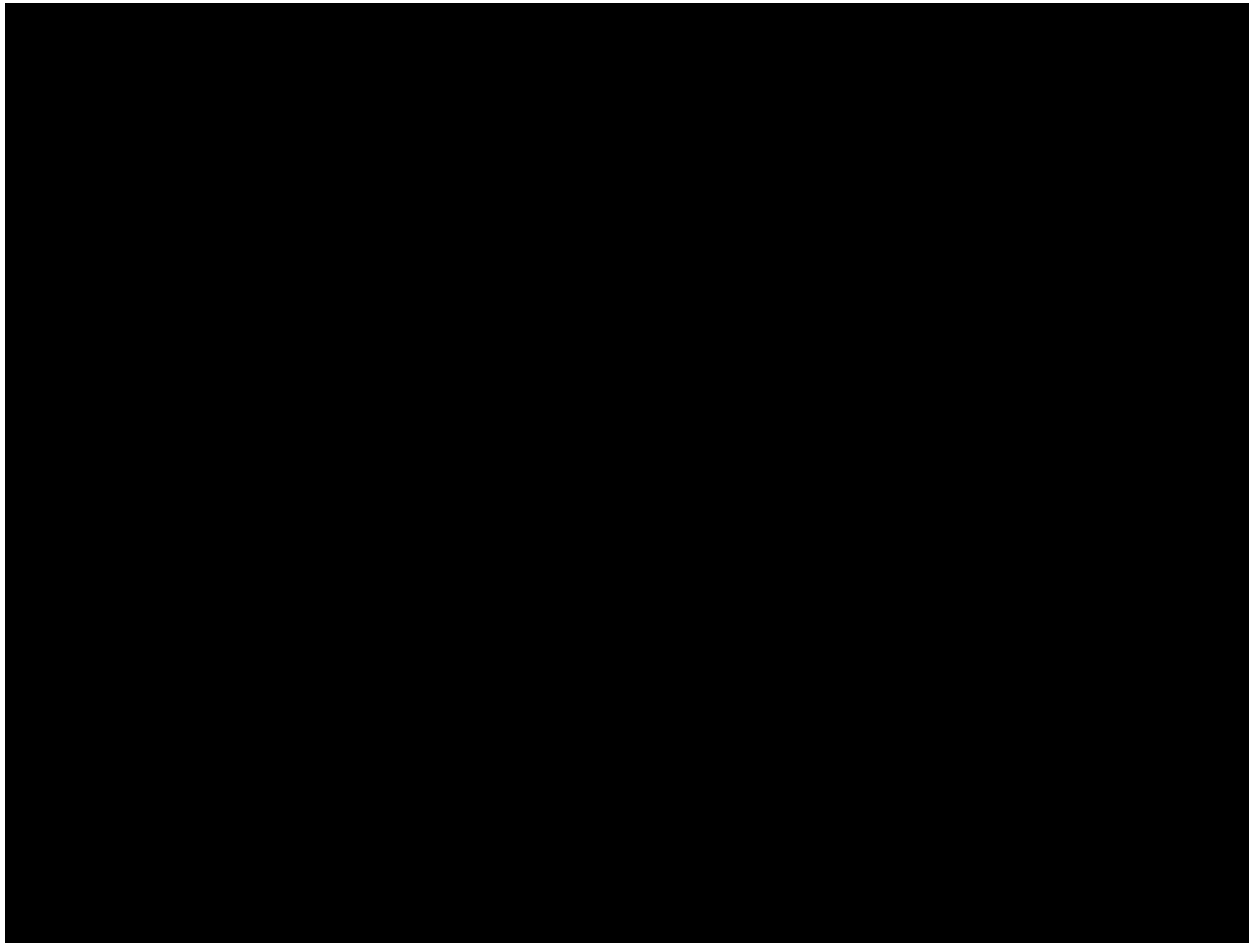
Worker (Gathers resources)

Base (Builds workers with resources)

Breaking the Game with AI and XAI

L2G: Learn to Gamebreak

- Players: Two AI agents complete against each other
- Convolution Neural Networks predict who is winning the game
- XAI Methods (SHAP and GPs) identify the winning advantage for the players
  – i.e., which feature contribute most to winning probability



Mission Engineering Insight with XAI:
Change the winning advantage found by XAI to shift the game balance

Summary and Conclusions

SE + M&S + XAI provide layers/views of information and tools for different target audiences to aid in design and evaluation of complex systems.

We highlighted ongoing research projects performed by the GMU's C4I & Cyber Center, focusing on technical achievements that addressed SE + M&S + XAI.

We applied state-of-the-art from knowledge representation and XAI to Systems, Systems-of-Systems and Mission Engineering problems (*e.g.*, Probabilistic Ontologies, MEBN, SHAP, XAI).

These capabilities are applicable to a wide spectrum of problems in IoT, Autnomous Vehicles, Cybernetics, Industry 4.0, Smart Cities, *etc*.

Finally, our presentation addresses and advances the tools and techniques required for System Engineering for AI (SE4AI).

References (Published papers on these projects)

- **Systems Level AI and XAI:**
  - Raz et., al., Test and Evaluation of Reinforcement Learning via Robustness Testing and Explainable AI for High-Speed Aerospace Vehicles, IEEE Aerospace Conference 2022

- **Systems of Systems Level AI and XAI:**
  - S. Matsumoto, A. Barreto, P. C. G. Costa, B. Benyo, M. Atighetchi and D. Javorsek, "Dynamic Explanation of Bayesian Networks with Abductive Bayes Factor Qualitative Propagation and Entropy-Based Qualitative Explanation," *2021 IEEE 24th International Conference on Information Fusion (FUSION)*, 2021, pp. 1-9, doi: 10.23919/FUSION49465.2021.9626961.

- **Mission Engineering Level AI and XAI**
  - Dachowicz, A., Mall, K., Balasubramani, P., Maheshwari, A., Raz, A. K., Panchal, J. H., & DeLaurentis, D. A. (2022). Mission Engineering and Design Using Real-Time Strategy Games: An Explainable AI Approach. *Journal of Mechanical Design*, *144*(2).
  - DeLaurentis, D. A., Panchal, J. H., Raz, A. K., Balasubramani, P., Maheshwari, A., Dachowicz, A., & Mall, K. (2021, August). Toward Automated Game Balance: A Systematic Engineering Design Approach. In *2021 IEEE Conference on Games (CoG)* (pp. 1-8). IEEE.

Unsolved Problems in ML Safety*

Robustness Create models that are resilient to adversaries, unusual situations, and Black Swan events.	**Alignment** Build models that represent and safely optimize hard-to-specify human values.
Monitoring Detect malicious use, monitor predictions, and discover unexpected model functionality.	**Systemic Safety** Use ML to address broader risks to how ML systems are handled, such as cyberattacks.

Systems Engineering for AI (SE4AI)

SE Test and Evaluation

- Input and output testing
- Sensitivity Analysis
- Design of Experiments

Explainable AI

- Algorithmic level
- Sub-system level
- System level

Verification and Validation

- Expected System Behavior
- Stakeholder Needs
- External system interface

Systems Engineering of AI is needed to help address these problems and transition AI into practical systems

Hendrycks, Dan, Nicholas Carlini, John Schulman, and Jacob Steinhardt. "Unsolved problems in ml safety." *arXiv preprint arXiv:2109.13916* (2021)
**Lewis, David K. 1973. *Counterfactuals*. Cambridge: Harvard University Press.

32nd Annual INCOSE
international symposium

hybrid event

Detroit, MI, USA
June 25 - 30, 2022

www.incose.org/symp2022