Test and Evaluation, Verification and Validation of Autonomous Systems

Safe and Secure Software and Systems

9 June 2015

Matthew Clark
V&V of Autonomous Systems
Autonomous Control Branch (AFRL/RQQA)
Air Force Research Laboratory
ASD/R&E framework of technical coordination groups

**Communities of Interest (COIs)**

**Missions**
- Cyber Security
- Counter IED
- Counter WMD

**Systems**
- Electronic Warfare
- Air Platforms
- Human Systems
- Info Systems
- Sensors & Processing
- Weapons Technologies
- Space
- Engineered Resilient Systems

**Enabling Science & Technology**
- Advanced Electronics
- Materials & Processes
- Energy & Power Technologies
- Autonomy
- ASBREM

Capabilities enabled by advanced technologies and systems

Integrating multiple technologies into complex systems

S&T with multiple applications
Hard Problems:

State-Space Explosion
- Algorithmic decision space is complex, adaptive, and cannot be exhaustively searched, examined, or tested

- Unpredictable Environments:
  - Autonomous systems operate in unknown, untested environmental conditions
  - Autonomous systems are difficult to assure correct behavior in a countless number of environmental conditions

- Emergent Behavior
  - Interactions between systems and system factors may generate unintended consequences
  - Autonomous systems are difficult to sufficiently capture and understand all intended and unintended consequences

- Human-Machine System
  - Handoff, communication, and interplay between operator and autonomy are key enablers for the trust and effectiveness of an autonomous system
Hard Problems:

State-Space Explosion
- Algorithmic decision space is complex, adaptive, and cannot be exhaustively searched, examined, or tested

- **Unpredictable Environments:**
  - Autonomous systems operate in unknown, untested environmental conditions
  - Autonomous systems are difficult to assure correct behavior in a countless number of environmental conditions

- **Emergent Behavior**
  - Interactions between systems and system factors may generate unintended consequences
  - Autonomous systems are difficult to sufficiently capture and understand all intended and unintended consequences

- **Human-Machine System**
  - Handoff, communication, and interplay between operator and autonomy are key enablers for the trust and effectiveness of an autonomous system
- **Emergent Behavior**
  - Interactions between systems and system factors may generate unintended consequences
  - Autonomous systems are difficult to sufficiently capture and understand all intended and unintended consequences

- **Human-Machine System**
  - Handoff, communication, and interplay between operator and autonomy are key enablers for the trust and effectiveness of an autonomous system
- Autonomous systems are difficult to sufficiently capture and understand all intended and unintended consequences

- **Human-Machine System**
  - Handoff, communication, and interplay between operator and autonomy are key enablers for the trust and effectiveness of an autonomous system
OSD Autonomy COI TEVV vs. AFRL Autonomy TEVV Taxonomy

- **Methods, Metrics, & Tools for Requirements Development and Analysis**
  - Precise, structured standards to automate requirement evaluation for testability, traceability, and de-confliction

- **Evidence based Design and Implementation**
  - Assurance of appropriate decisions with traceable evidence at every level of design to reduce the current T&E burden

- **Cumulative Evidence through RDT&E, DT & OT**
  - Progressive sequential modeling, simulation, test and evaluation

- **Run time behavior prediction and recovery**
  - Real time monitoring, just-in-time prediction, and mitigation of undesired decisions and behaviors

- **Assurance Arguments for Autonomous Systems**
  - Reusable Assurance Case based on previously evidenced building blocks
CMU Study on Software Cost

Despite Autonomy!


**Major cost savings through rework avoidance by early discovery and correction**

A $10k architecture phase correction saves $3M

**Rework and certification is 70% of SW cost, and SW is 70% of system cost.**

**Sources:**


Requirements Development & Analysis

Precise, structured standards to automate requirement evaluation for testability, traceability, and de-confliction

RESEARCH OPPORTUNITIES:
• Formal expressibility of natural language
• Dynamic requirements analysis and feedback
• Semantics for non-functional, high-level requirements
• Formalized probability requirements
• Automatic Test Case and Assurance Case structures
• Dynamic autonomy based requirements generation (Learning Algorithms)
Evidence Generation During Design

Guarantee appropriate decisions with traceable evidence

RESEARCH OPPORTUNITIES:

• Scalable, unifying formal, mathematical, construct across design abstractions / design methods / test results
• Trust / transparency in design
• Link architecture to “correct by construction” synthesis
• Incremental analysis and proof reuse
• Learning based formal model adaptation
• Scalable analysis methods of formal models for reliability, probability, timing, & hybrid systems
Cumulative Evidence Through Research, Developmental, and Operational Test

Progressive sequential modeling, simulation, test and evaluation

RESEARCH OPPORTUNITIES:
• Design of Experiments (DOE) methods for non-deterministic / learning algorithms
• Transparency / Instrumentation of System Under Test (SUT) Learning Algorithms
• Leveraging formal models in test isolation / design updates / test reuse

TRMC VIDEO
https://www.youtube.com/watch?v=HSvDEbqjmCA&feature=youtu.be
Enable Real time monitoring, just-in-time prediction, and mitigation of undesired decisions and behaviors

RESEARCH OPPORTUNITIES:
• Run time analysis and scalable prediction of undesired constraints
• Dynamic switching in the incorporation of certifiable recovery methods – ACAT, SAA, Geo-Fence etc.
• Learning based auto constraint generation and transparency
• Real time computation of stochastic hybrid systems
• Run time constraints derived from assume-guarantee based designs
Compositional Case Generation

Enable reusable evidence building blocks

RESEARCH OPPORTUNITIES:

• Formal mathematical construct for argument based notations like Goal Structuring Notation (GSN)
• Formal analysis of structures and semantics of arguments linked to requirements
• Operational Research (OR) based argument generation
• Formal mapping of implicit regulations to explicit arguments for autonomy
• Learning based argument generation
• Formal argument composition and analysis
We’re Making Progress.. but

A lot of Work in Making Better – More Verifiable Designs
We’re Making Progress.. but

A lot of Work in Making Better – More Verifiable Designs

We are starting to change how we Formally and Consistently Argue Those Designs
We’re Making Progress.. but

A lot of Work in Making Better – More Verifiable Designs

We are starting to change how we Formally and Consistently Argue Those Designs

T&E Community is beginning to Adopt New Verification Methods
We’re Making Progress.. but

We are starting to change how we Formally and Consistently Argue Those Designs

A lot of Work in Making Better – More Verifiable Designs

T&E Community is beginning to Adopt New Verification Methods

Formal Designs Still suffer from Scalability / Interoperability
We’re Making Progress.. but

Sys Engineering Community needs **New (Formal Tools)** For Requirements

A lot of Work in Making Better – More Verifiable Designs

We are starting to change how we **Formally and Consistently Argue** Those Designs

T&E Community is beginning to **Adopt New Verification Methods**

Formal Designs Still suffer from **Scalability / Interoperability**