Testing and Evaluation of Cognitively Advanced Autonomy

Autonomous and Intelligent Systems Division
Applied Research Lab – Penn State University

Dr. John Sustersic
Research Associate and Head (acting)
Autonomy, Perception and Cognition Department
Autonomous and Intelligent Systems Division
Applied Research Laboratory
Office: 814-863-3015
Email: jps263@arl.psu.edu

High Assurance Human-Machine Systems

- How does Airline Transportation achieve exceptional (though still imperfect) performance?
  - Semi-Controlled Environment
    - Airspace rules designed to provide safe operating environment.
    - Sufficient weather sensing and prediction to reasonably ensure safe operations.
  - Certification of electro-mechanical systems.
    - Establish limits of safe operating envelope.
    - Expensive, necessarily slow to adapt to new technologies.
  - Qualification of Humans in system operation.
    - Demonstrated understanding of fundamental tasks, rules of the operating environment.
    - Demonstrated understanding of automatic systems to understand what they do, don't do, and when states change.

Tomorrow's Autonomous Systems

- Human-Machine Interaction and Collaboration
  - Expect similar 'rules' as today's manned systems.
- Interact with human collaborators using naturally (via language, gestures, illustrations, etc.)
  - Necessary for efficiency and practicality.
  - Semantics highly contextualized, often non-verbalized in normal interactions.

Example: "Go to the store and get some milk."

Which store? What kind of milk? How much milk? What do you do if the store is out of fat-free gallons of milk you normally buy?

Latent contextual semantics and the ability to adapt semantics naturally are essential for effective autonomy

Autonomy vs. Automatic

- An autonomous system is one in which decision are made in the complexity of environment-system interactions are intractable.
  - E.g. a person walking around Stockholm Central Station trying to find a conference.
- An automatic system exhibits known behaviors within a well-defined operating envelope.
  - E.g. elevator/lift, cruise control.

Tomorrow's Autonomous Systems in Human-Machine Collaboration

- Will the same approaches to achieve today's high reliability systems get us where we need to go?
  - Semi-Controlled Environment
    - Hard to imagine social changes required to sufficiently enable robotic/autonomous systems to participate.
  - Certification of mechanical systems.
    - Economic viability of pervasive certification questionable.
  - Qualification of Humans in system operation.
    - Those who might most benefit from autonomous aides might be less able to learn complex systems (e.g. elderly, sick, disabled, etc.)

Smart Phone evolution demonstrates commercial viability of technologies with intuitive interfaces and practical applications.

Of course, so does evolution of drone technologies; regulating agencies are struggling to keep up.

PennState Applied Research Laboratory

The Reality of Testing and Evaluating Autonomy

We must choose between:

Sufficiently restricted use cases s.t. sufficient T&E certification is affordable

↔

Constrained use cases s.t. hybrid T&E certification and performance-based qualification is acceptable

- Automatic vs. autonomous systems.
- To allow for more complexity in autonomy, must trade off ability to statically verify properties of the system.
  - Objective then becomes to bound system behavior, work to ensure system degrades/fails gracefully.

Understanding the Limits of Verification

- Consider a pencil standing vertically on the pointed end:
  - What happens?
  - Which way does it fall?
  - What can we predict?

Verification shows convergence of statistical model in limit

Verification shows abstract ‘understanding’ that the pencil will, fall; exactly how depends…

If we force a homogeneous solution to this prediction problem, we at best can have a stochastic description of expected outcomes.

By allowing for abstraction and heterogeneous models, we can express both verifiable components of the model and components that are not well suited for robust verification.

Understanding the Limits of Verification

- Consider a pencil standing vertically on the pointed end:
  - What happens?
  - Which way does it fall?
  - What can we predict?

How would we measure these things in the real world?

Even if we could, methods would not work for non-deterministic systems.

This is a deterministic, chaotic system. Can we not just measure more closely the precise initial conditions (vertical alignment, mass distribution, shape of tip, etc.) and use this information to determine outcome?

No. Other factors (air currents, vibrations in surface, etc.) can effect the outcome. Heisenberg uncertainty precludes limit cases.

Information Theory and T&E

- Occam's Razor: simpler hypotheses that explain data equally well are preferable because they are more testable.
- Minimum Description Length: formalizes Occam's Razor, best hypothesis gives best compression.
  - E.g. Consider Newton's Second Law:

$$\boldsymbol{F} = m_0 \boldsymbol{a}$$

Classical Mechanics

$$\boldsymbol{F} = \gamma(\boldsymbol{v}) m_0 \boldsymbol{a}_\perp + \gamma(\boldsymbol{v})^3 m_0 \boldsymbol{a}_\parallel$$

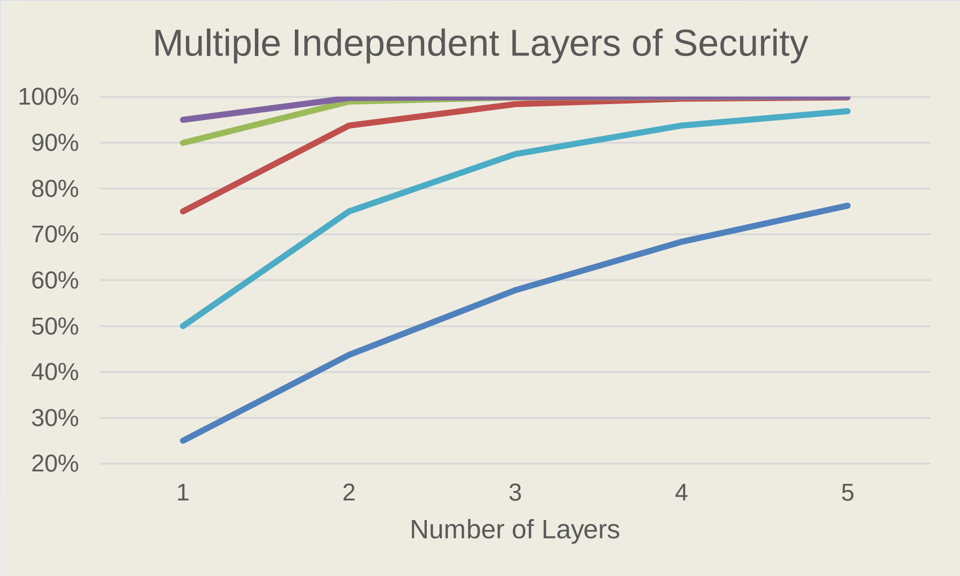
Relativistic Mechanics

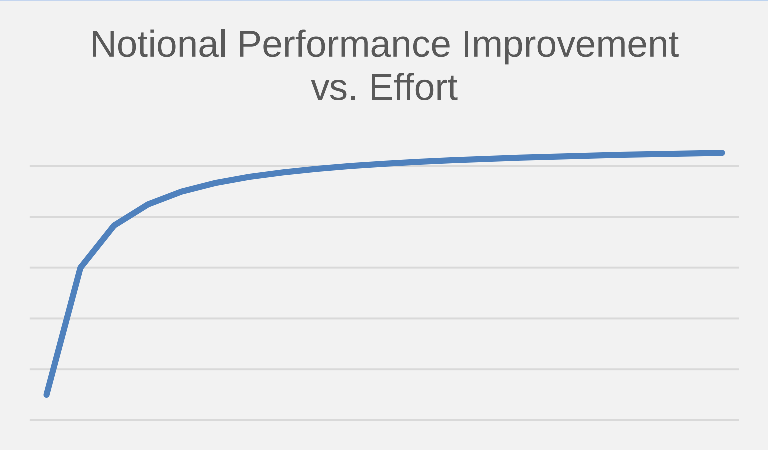
Practical T&E

- What are T&E ramifications for an AxV?
  - Simplifies perception: does not require ultra-precise measurements of many parameters to operate.
  - Bounds verification space: only need to verify that system can manage any *possible* eventuality.
    - Can relax this somewhat to show coverage for most ***likely*** states.

Taking a Cue from Information Assurance

- Perfect security is impossible.
- Best possible security is impractical.
- More cost-effective to layer multiple, independent security measures.





MILA

How can complex manned autonomous systems robustly and reliably perform in dynamic, unconstrained environments despite:

- Imperfect decisions
- Uncertain, incomplete, imprecise and contradictory data
- Equipment and decision failures?

- Multiple, Independent Layers of Assurance enables:
  - Mission-level constrained, bounded learning.
  - Intrinsic error and fault tolerance.
  - Rapid deployment of capabilities.
    - Less time wasted trying for incremental improvements at high costs in time and money.
  - High-reliable, robust autonomy.
  - Supports handling need-to-know and software assurances best practices.
- MILA requires non-bypassability, statistically independent layers

PennState Applied Research Laboratory

Theory Constrains Practical Intelligence

Just like human learning and intelligence:

Mathematical logic, computability theory, complexity theory, and learning theory imply that practical learning cannot be perfect.

1. Perfect learning is impossible.
2. Best possible learning is impractical.
3. Therefore, practical learning is about finding good heuristics.

Therefore:

Satisficing is a theoretical necessity in practical learning.

Teaching is needed, and allows for the transfer of knowledge that was gained at significant expense in money, time, blood, etc.

Summary

- Evolution of T&E practices needed to allow for more rapid integration of new technologies into AxVs.
  - Stay ahead of society risk tolerance.
- MILA approach likely most cost effective, agile enough to keep pace with technological advances.
- Bounded learning as a T&E solution, not headache.
- Certification + qualification for complex autonomy VV&A / T&E.