NAVAIR Autonomy TEVV Study

David Scheidt
dscheidt@weathergagetech.com

Motivation

NAVAIR requires a standing, in house, autonomy test, evaluation, verification and validation (ATEVV) capability that is capable of providing the necessary assurance argument required to field autonomous unmanned air vehicles.

Bootstrapping this capability requires that NAVAIR:

- Develop policies and practices for ATEVV
- Acquire a suite of tools suitable for ATEVV
- Establish skills and expertise in the use of these tools for autonomous system testing

There is a strong desire to efficiently accomplish this goal through strong cooperation with the larger ATEVV community.

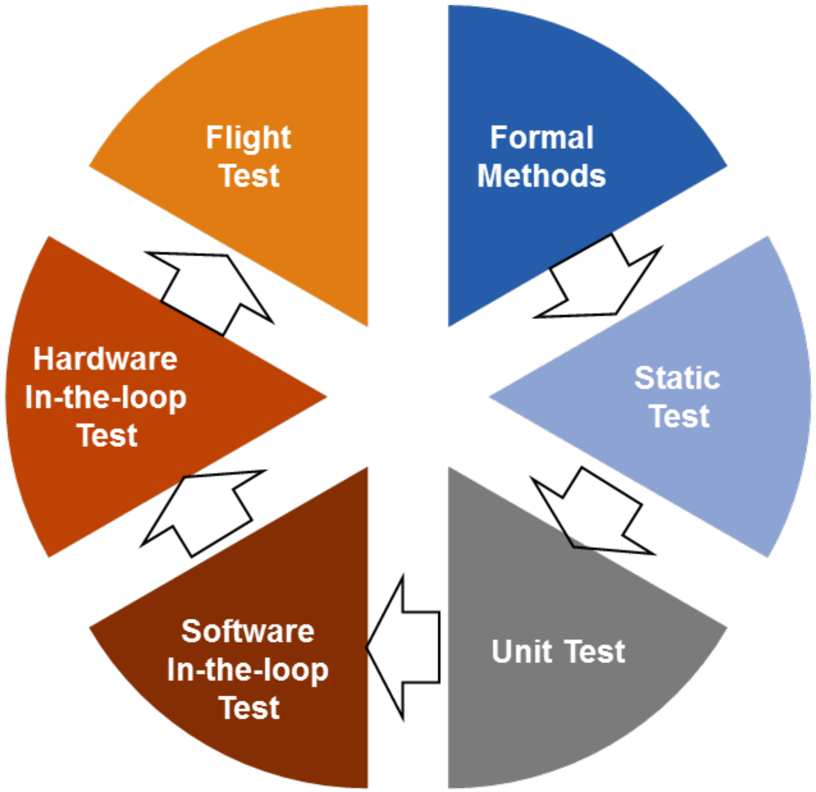
The MQ-8C Fire Scout prepares to land on the guided-missile destroyer, USS Jason Dunham (DDG 109), for the first time Dec. 16, off the Virginia coast. (Photo by Northrop Grumman)

Four key points from prior work

1. Defining metrics on “levels of autonomy”, is not useful and consumes resources. What is important is to develop the tools and methods for testing these systems. – DSB 2012/DSB 2015.
2. Autonomy TEVV is not just a DT&E/OT&E problem. TEVV must be involved from requirements analysis through deployment. – OSD ATEVV Investment Strategy
3. No single method/tool is sufficient to produce an assurance argument. A complex, integrated process that utilizes multiple tools is required – Autonomy T&E Infrastructure Gap Study, 2016
4. TEVV of autonomous systems is further complicated by the complex, integrated nature of the “thinking” part of an autonomous system. Validating that each algorithm performs to spec is insufficient.

Weather Gage Technologies, LLC
dscheidt@weathergagetech.com

A Simple Look at the ATEVV Process



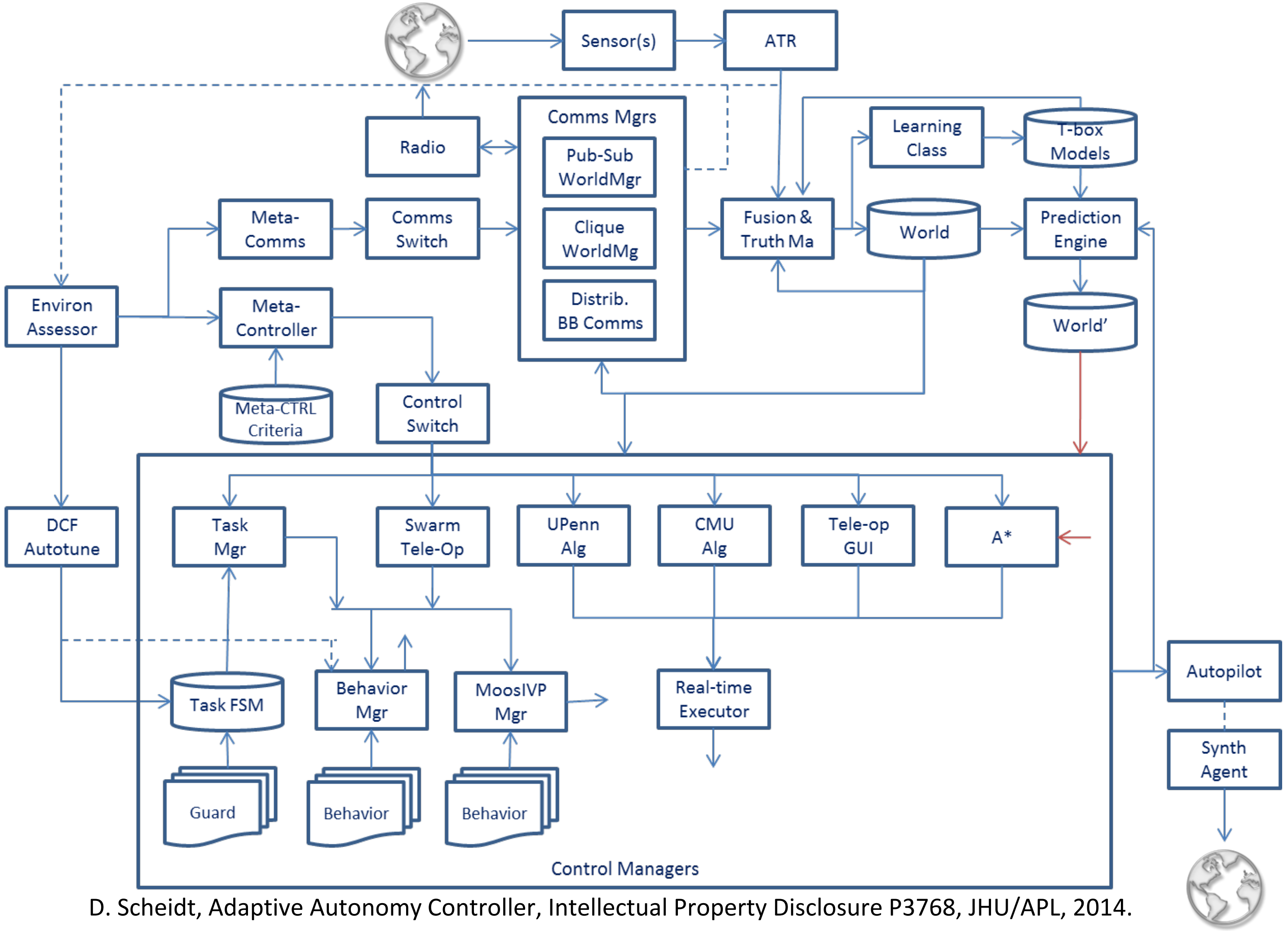
Torens, C., Adolf, F. (2014), "V&V of Automated mission planning for UAS", *NATO SCI-274 Workshop Verification and Validation of Autonomous Systems*, Imperial College, London, June 24-25.

Sample Autonomy Architecture

Sample Autonomy Architecture

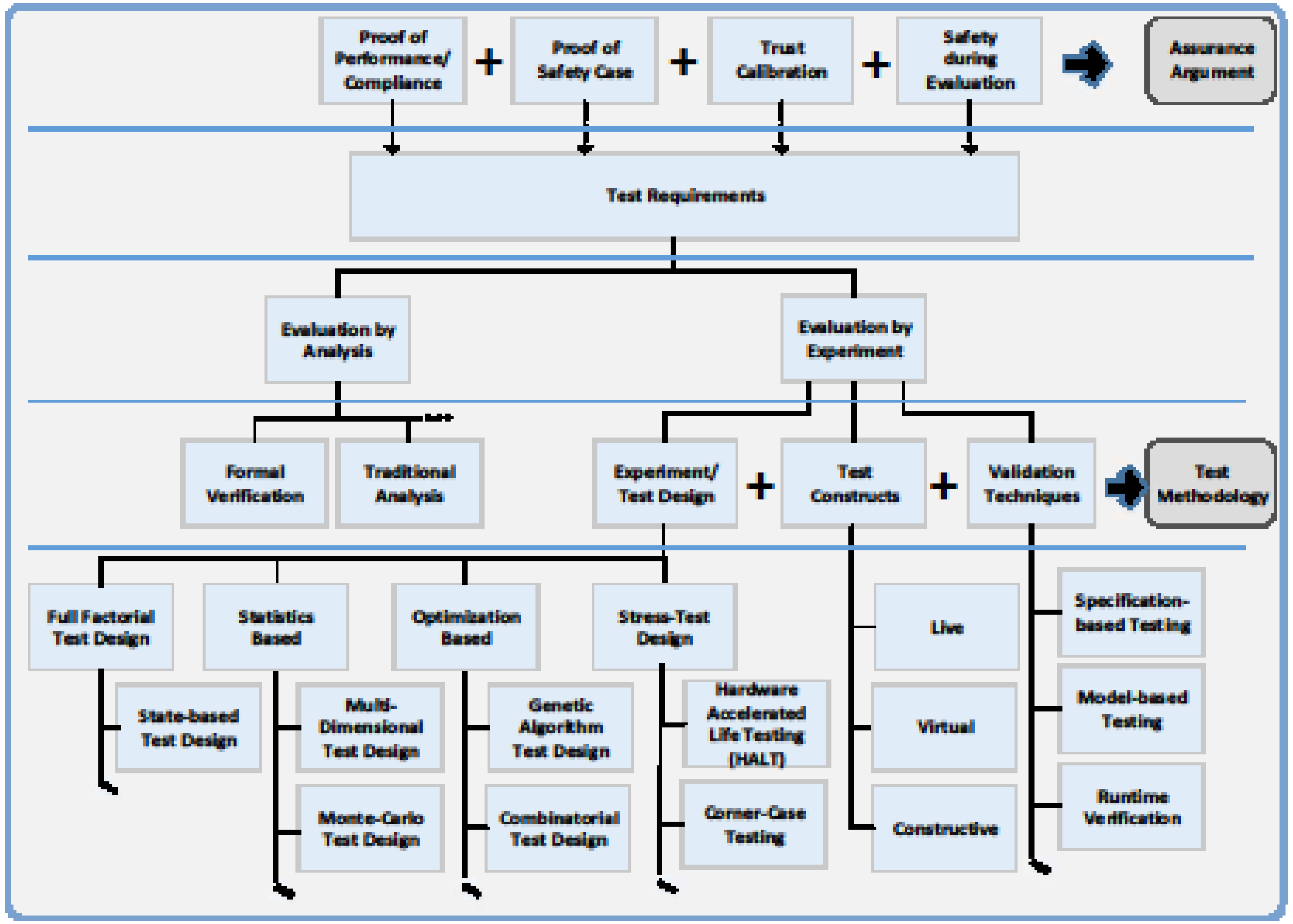
- Includes > 20 distinct reasoning algorithms
- Diverse Classes of Reasoning Algorithms
  - Neural Networks
  - Production-rule Systems
  - Search Algorithms
  - Swarm/Potential Field
  - Bayesian Belief Nets

Independent validation of these algorithms is insufficient. An assurance argument that guarantees that the properties that emerge from the interactions between the algorithms are satisfactory is necessary.



D. Scheidt, Adaptive Autonomy Controller, Intellectual Property Disclosure P3768, JHU/APL, 2014.

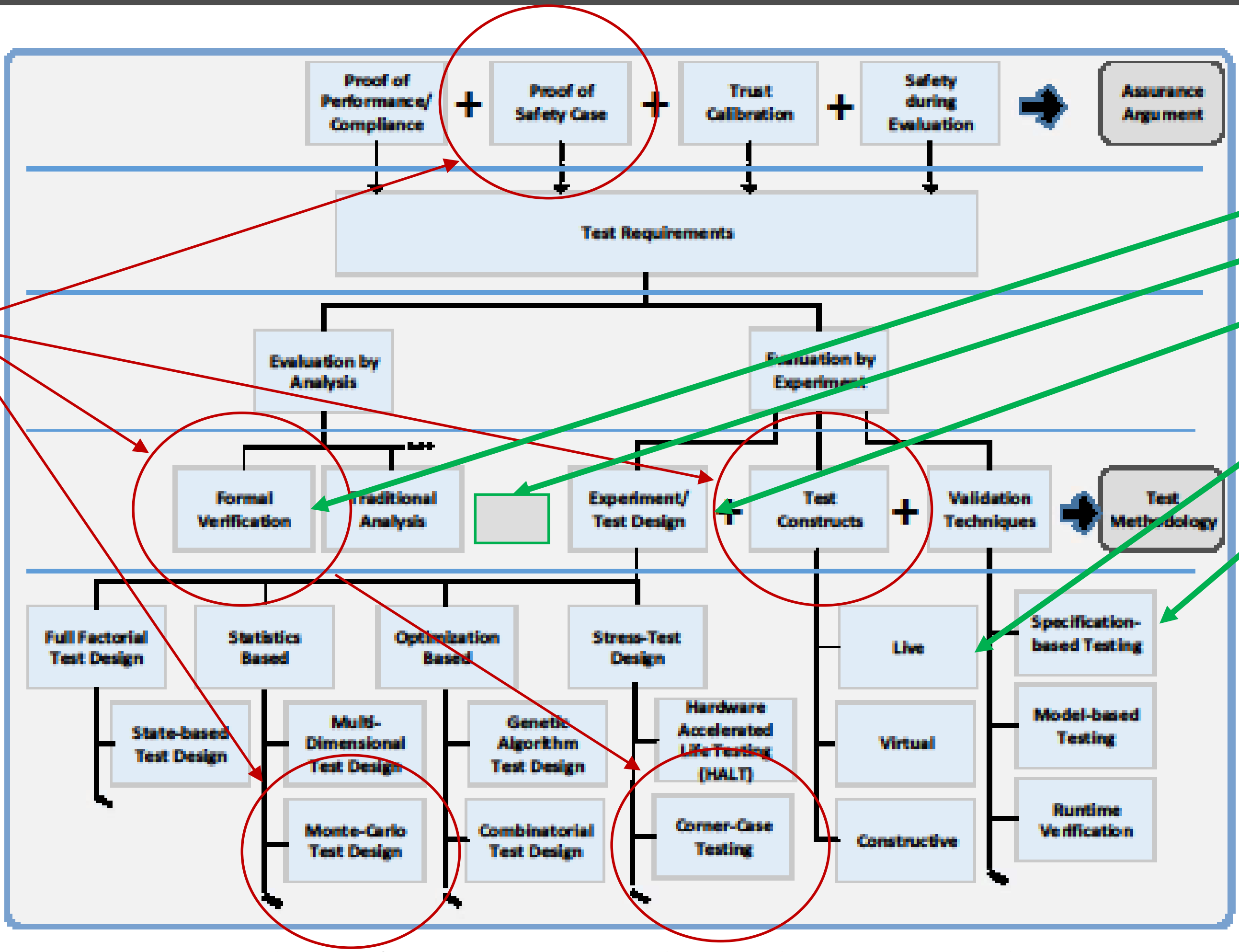
A More Expressive Look at TEVV Processes and Tools



Patel, Rohintan, Autonomy Test and Evaluation Infrastructure Gap Study – Phase I Final Report, Georgia Tech Research institute, Project D7533, Jan. 2016.

Key Contribution of This Study

How do these relate to each other? The answer is that we don't really know, and we're not really likely to know until somebody (NAVAIR) attempts an end-to-end ATEVV.



Different tools address different phases:

Keymera
RIOT
ATAS
TACE/WFN
RAPT

Patel, Rohintan, Autonomy Test and Evaluation Infrastructure Gap Study – Phase I Final Report, Georgia Tech Research institute, Project D7533, Jan. 2016.

Study Process

An 8 step process to develop the Autonomy TEVV process is planned:

- Scope the problem
- Define the requirements
- Define the metrics, units of measure
- Cross matrix the scope and the requirements
- Identify and access tools and methods
- Define assurance process
- Map existing tools to the process and identify gaps
- Go experimenting

dscheidt@weathergagetech.com

Scope the problem

For a candidate system, which is likely to be either an unmanned air system (UAS) conducting Intelligence, surveillance and reconnaissance; logistics UAS

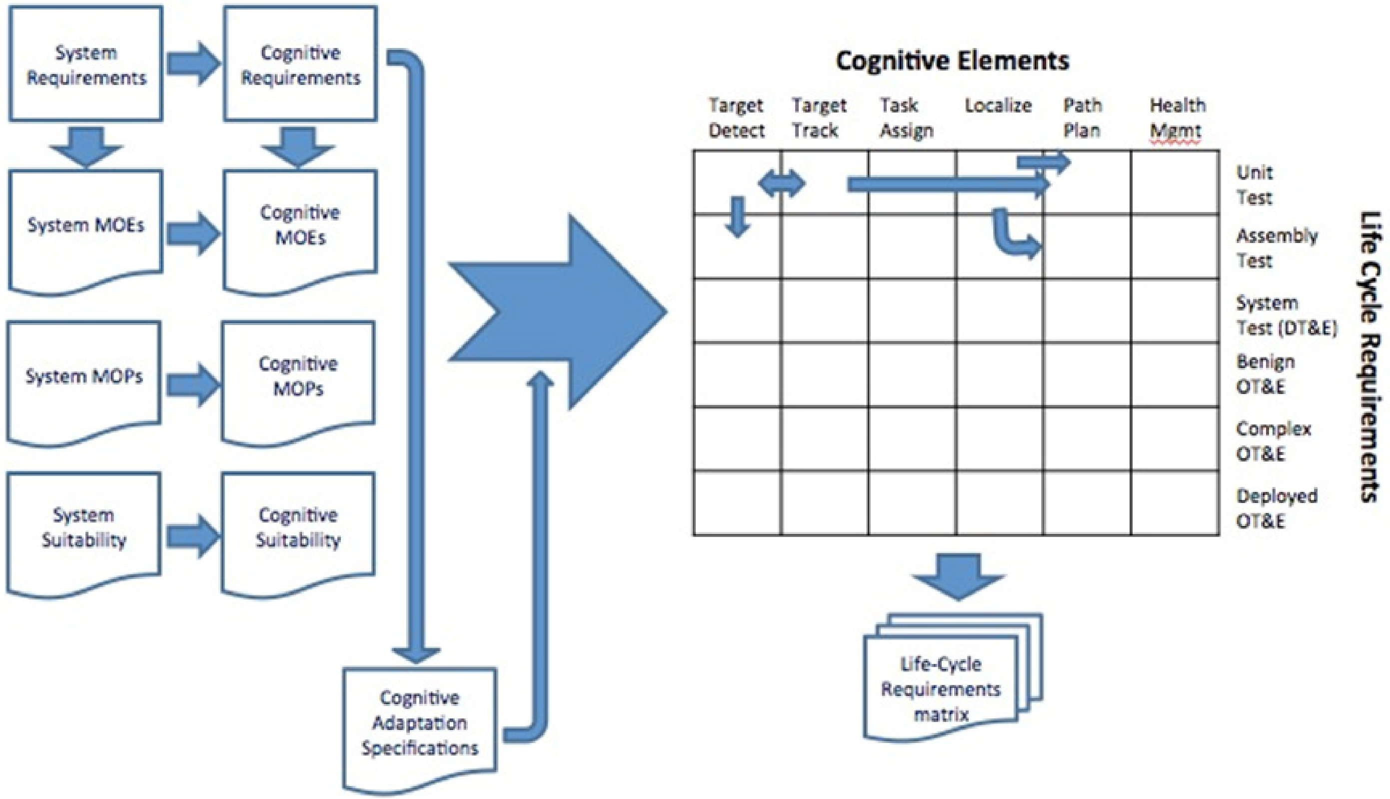
- What is the “autonomous” about the system?
  - What kinds of decisions are made by the autonomous system: perception, localization, path planning, learning of target classification, learning target behaviors, learned control policies and so on
- What is the operating environment like?
  - Static or Dynamic
  - Simple or complex
  - Includes other decision-makers, are they friendly, neutral or adversarial?
- What are the possible interactions between the autonomy and the outside world?

Anticipated insight and product – Identifying gaps in terms, ontologies and languages that are required to define operational domain and conditions.

Define Cognitive Element TEVV Requirements

Our focus is on exploring the ATEVV process, as such, our requirements must fully exercise the ATEVV process, yet may be incomplete.

As with all well-formed requirements, they must be testable.



D. Sparrow, et. al., Appendix C - Requirements and Metrics, A Framework for Evidence-based Licensure of Adaptive Autonomous Systems, *Institute for Defense Analysis Paper P-5325, Mar, 2016.*

Requirements Definition – Step 1

System Specifications

Representative System Specification Requirements for Autonomous Tractor Trailer	***Associated Metrics***
Measures of Performance	
"stops at red lights"	Brakes successfully activate to stop vehicle within *n* meters of red light or associated striping
"follows local traffic laws."	
Measures of Efficiency	t_0 = start time t_p = predicted arrival time t_a = actual arrival time
Arrives "on-time"	*on-time metric* = $\Delta(t_p - t_a)$
Arrives "as fast as possible"	*as_fast_as_possible* = min t_a (t_a- t_0)

The first product of the autonomy requirements process is a detailed specification that defines requirements for the entire system, which include the physical plant as well as the autonomous decision-making apparatus.

D. Sparrow, et. al., Appendix C - Requirements and Metrics, A Framework for Evidence-based Licensure of Adaptive Autonomous Systems, *Institute for Defense Analysis Paper P-5325, Mar, 2016.*

Requirements Definition – Step 2

The second product of the autonomy requirements process is the development of the cognitive specifications which separates out the requirements that are specific to the decision-making process.

Cognitive Specifications

Representative Cognitive Requirements for Autonomous Tractor Trailer	***Associated Metrics***
Must identify suitable route	
Must decide to stop at stop signs	Makes the decision to stop at the stop sign
Must decide to continue on from stop sign iff way is clear	
Must decide to continue on from stop sign iff	

D. Sparrow, et. al., Appendix C - Requirements and Metrics, A Framework for Evidence-based Licensure of Adaptive Autonomous Systems, *Institute for Defense Analysis Paper P-5325, Mar, 2016.*

Requirements Definition – Step 3

Adaptability Decomposition Matrix

Autonomous Tractor Trailer Adaptability Decomposition	**Static Uncertainty**	**Dynamic Uncertainty**	**Evolutionary Uncertainty**
ID Suitable Route	Localization Error Road Closures Road Conditions	Traffic (other drivers)	Traffic patterns
Stop at Stop Signs	Lighting Conditions (ability to sense) [illegible]	Other drivers	Rome drivers ≠ Columbus drivers

The third product of an autonomous system requirements analysis is the Adaptability Matrix, which identifies change that must be managed by the autonomous system.

D. Sparrow, et. al., Appendix C - Requirements and Metrics, A Framework for Evidence-based Licensure of Adaptive Autonomous Systems, *Institute for Defense Analysis Paper P-5325, Mar, 2016.*

Requirements Definition – Step 4

Cognitive Decomposition Matrix

Autonomous Tractor Trailer Cognitive Decomposition	**Localization**	**Path Planning**	**Path Assessment**	**Object Detection**	**Object Classification**	**Object Tracking**	**Fault Management Diagnosis**
Must ID suitable route	Localization Orientation	Select route along allowed roads	Assess timing of route	Learn of historical blockages			
Must reroute when necessary	Localization Orientation	Select route along allowed roads	Assess timing of route	Accept communicated blockages			
Must stop at stop signs	Localize self Localize sign Locacalize stop stripe	Route to location at strips, or behind car in front		- Cars (in front) - Cars @ other signs - Cars or roads not impeded by sign (e.g. on primary	Cars	Cars - Cars	

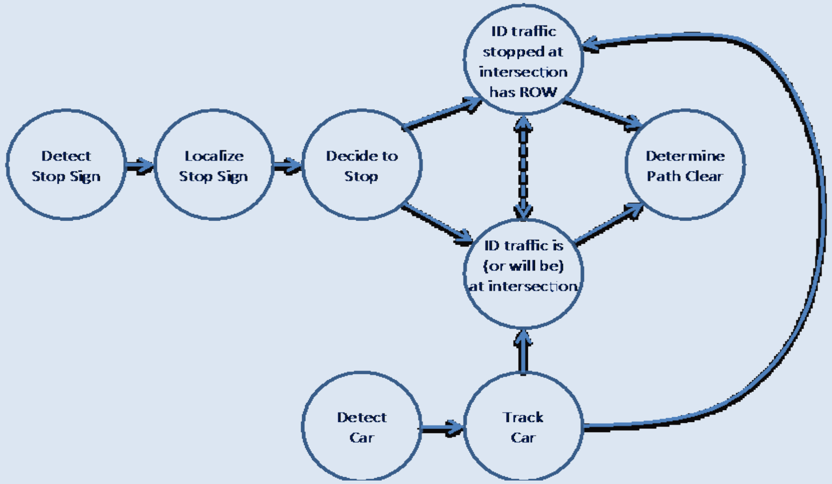
The forth product of the autonomy requirements process is the Cognitive Decomposition which breaks down decision-making requirements into sub-requirements for each component within the cognitive architecture.

D. Sparrow, et. al., Appendix C - Requirements and Metrics, A Framework for Evidence-based Licensure of Adaptive Autonomous Systems, *Institute for Defense Analysis Paper P-5325, Mar, 2016.*

Requirements Definition – Step 5

The final product of the requirements process is cognitive process model that illustrates the relationships between cognitive components and a dependency matrix that enumerates the required quality of cognitive inputs and the produced quality of cognitive outputs. By explicitly enumerating cognitive component dependencies we may produce assurance traceability from disparate sources.

Cognitive Dependency Mapping



Cognitive Component	Input Requirements	Output products
Detect Stop Sign	--	P_D > 99.999 @ 50m range
Localize Stop Sign	Detection 99.999 @ 40m range Own-vehicle localization (error < 1m^2)	Stop sign position (error < 1 m^3) @ 50m range

D. Sparrow, et. al., Appendix C - Requirements and Metrics, A Framework for Evidence-based Licensure of Adaptive Autonomous Systems, *Institute for Defense Analysis Paper P-5325, Mar, 2016.*

Identify and characterize available TEVV tools

For each tool/method identify –

- Scope of the tool, what decisions it can be used to test, class of vehicle
- Assumptions/inputs required by the tool
- Products/outputs by the tool and the assurance/invariants represented by those products

Weather Gage Technologies, LLC
dscheidt@weathergagetech.com

Defining an Assurance Process

After defining the requirements in detail, and identifying suitable ATEVV tools, an Assurance Plan defines the assurance methods that will be used to engender trust of each cognitive component.

Assurance Plan

Autonomous Tractor Trailer Test and Evaluation Decomposition	**Formal Methods**	**Static Test**	**Unit Test**	**Software in the loop**	**Hardware in the loop**	**Flight Test**
Car detection			Google car ride-along testing			
Localization			Assess timing of route			GPS evidence
Path Planning	Theorem proving to show no deadlock	Hoare logic proof		-		
Tire pressure						UL tested

D. Sparrow, et. al., Appendix C - Requirements and Metrics, A Framework for Evidence-based Licensure of Adaptive Autonomous Systems, *Institute for Defense Analysis Paper P-5325, Mar, 2016.*

Map existing tools to the process and identify gaps

Select existing tools and techniques for use in ATEVV for our target system. Define a subset of requirements we expect to validate (note that a full validation is likely to be prohibitively expensive).

From the sets of tools identified produce a 4D mapping of Scope x Requirements x Cognitive Element x Tool/method in which each cell is (roughly) formulated as a Hoare triplet;

{Preconditions} <method> {Postcondition}

This provides us with a formalism that can be used to combine TEVV product results into an assurance argument. Note that it is more important to create and end-to-end chain that spans the process than it is to comprehensively address all requirements.

Products from this stage include:

- Development of a cross-TEVV phase assurance formalism.
- Identification of gaps.
  - What required input to tools/methods is not readily available?
  - What conceivable process/tool could address that gap?
  - Can the products be combined to produce a complete assurance argument? If not, what are those gaps?

dscheidt@weathergagetech.com

Go experimenting

Conduct a deep experimentation program in which all phases of the ATEVV process are used to produce a ***limited*** assurance argument.

- Did the tools produce the expected products? If not, why?
- Were the products sufficient to support the argument envisioned? If not, why?

Consider Policy Issues

Based upon our experimental results:

- What are the limits of ATEVV as we now understand them?
- What are the ethical and legal ramifications associated with our ability to understand and assure the performance of autonomous systems?

- What should a policy that appropriately divides the authority and responsibility for autonomous unmanned vehicle operations between:
  - Operator
  - Commander
  - Peers (i.e., combatants w/o control over UAS)
  - Acquisition Community
  - Test Community
  - Development Contractor

Weather Gage Technologies, LLC
dscheidt@weathergagetech.com

Acknowledgement - This work is supported by NAVSEA under Contract N00024-13-D-6400.

Weather Gage Technologies, LLC
dscheidt@weathergagetech.com