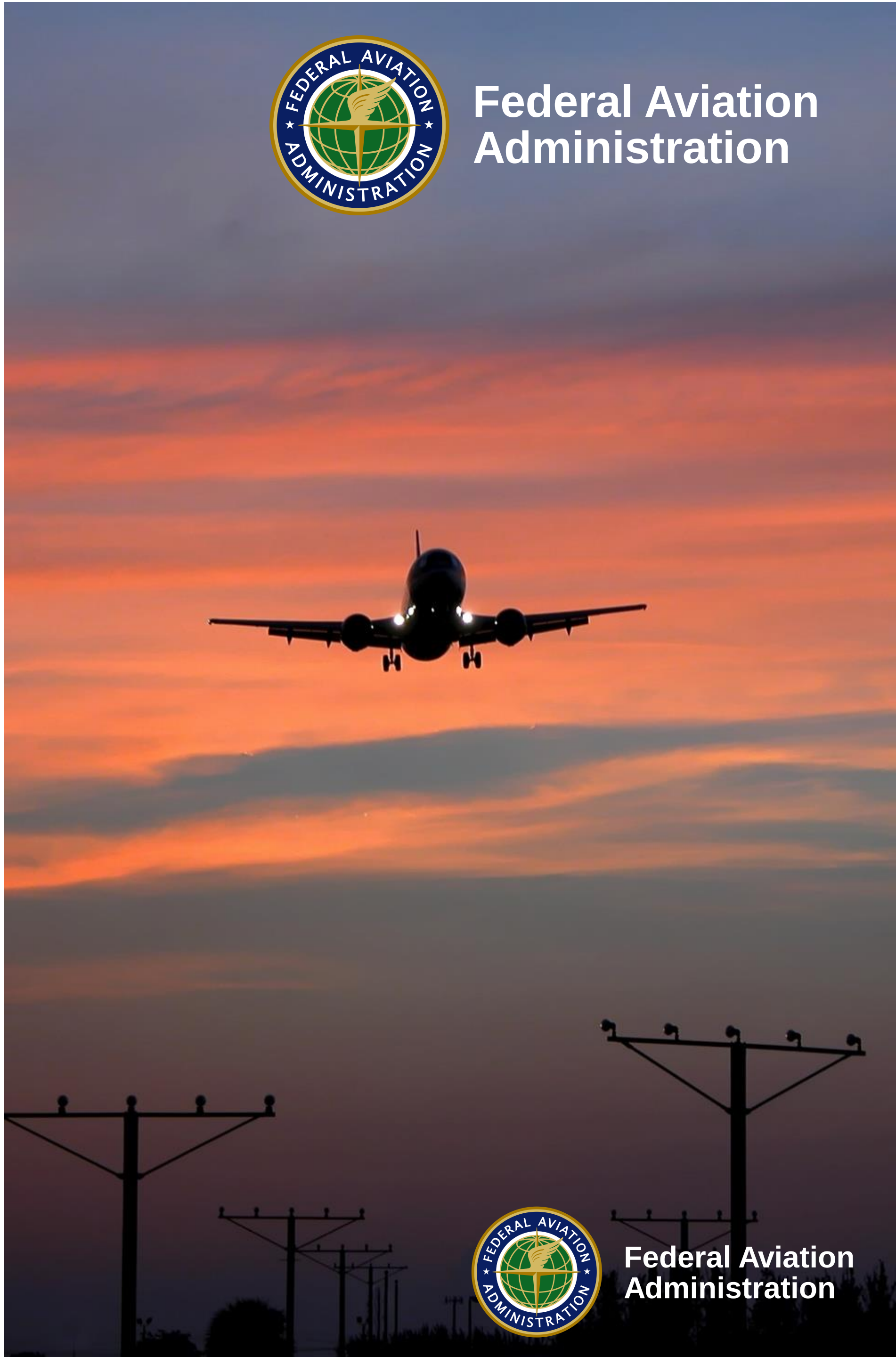
Artificial Intelligence and Certification



George Romanski,

Chief Scientific and Technical Advisor
Aircraft Computer Software

Date: May, 2019

Challenges of AI/ML Certification for Avionics Systems

- Artificial Intelligence and Machine Learning approaches have enjoyed much success
- Can they be trusted in safety critical situations
- Deployment is pushing the boundaries of innovation
- Approval by authorities appears to be lagging
- New approaches are being explored

AI has existed for a while

- Expert System, Artificial Intelligence was a HOT-TOPIC in THE ‘80’s
- They were mostly Inference Engines based on programming languages
  - LISP
  - PROLOG, etc.
- They were hard to program and limited by computing power

Artificial Neural Networks

- New paradigm evolved over last 10 years
- Incredible growth of computing power
- Huge volumes of data available cheaply
- New approaches mimicking operations of brains (sort of)

Computing power spurt

- Game computers demand more realism
  - Ray tracing are used to draw more realism into graphics
    - This requires huge multiply-add operations on arrays of data values
    - High speed required to repeat operations in video frame speeds
    - Co-processors developed to handle simple computations
    - Video Cards developed with multiple processor cores, or vector processing e.g. NVIDIA
  - Tighter memory/processor coupling
    - Instruction/Data cacheing

Data Availability

- Big Data – through database scraping
  - Data storage became “cheap”
  - More transactions through higher throughput on Internet
  - Data stored “in the cloud”
- Systems can “learn” from historical data
- This was exploited by “deep pockets”
  - Amazon – shopping cart suggestions
  - Google – Search engines
  - Facebook – Social-media linking

Automation based on Artificial Intelligence

There are many kinds of AI approaches, and many new ones are being invented

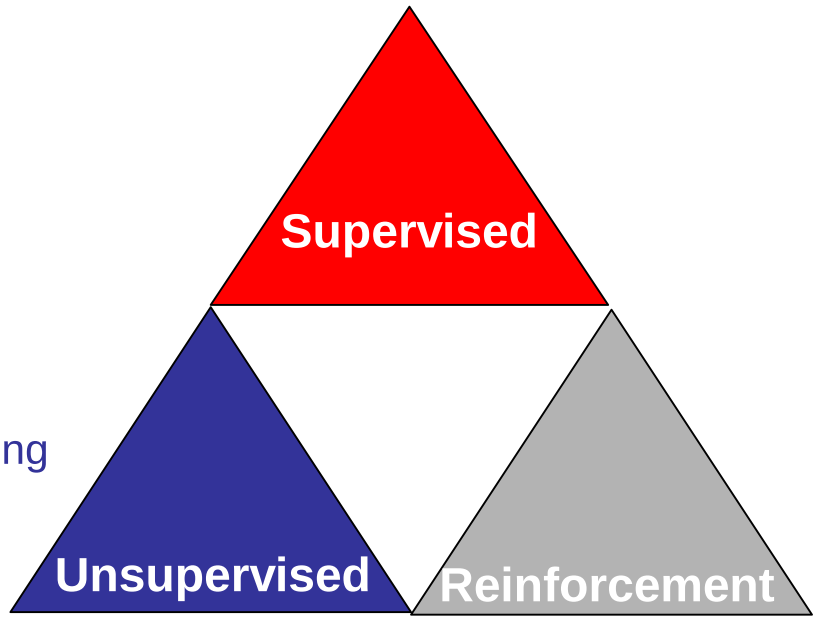
- Rule Based, Behavior trees, State machines
- Neural Networks –
  - Unsupervised
    - Learning by Data clustering
  - Supervised
    - Labeled Data
  - Reinforcement Learning
    - Heuristic reward function to extrapolate information

prominent due to increase in computing resources

Learning types

- Labeled Data
- Direct Feedback
- Predict outcome/future

Image Recognition



Data Clustering

- No labels
- No feedback
- “Find hidden structure”

Control Systems

- Decision process
- Reward System
- Learn series of actions

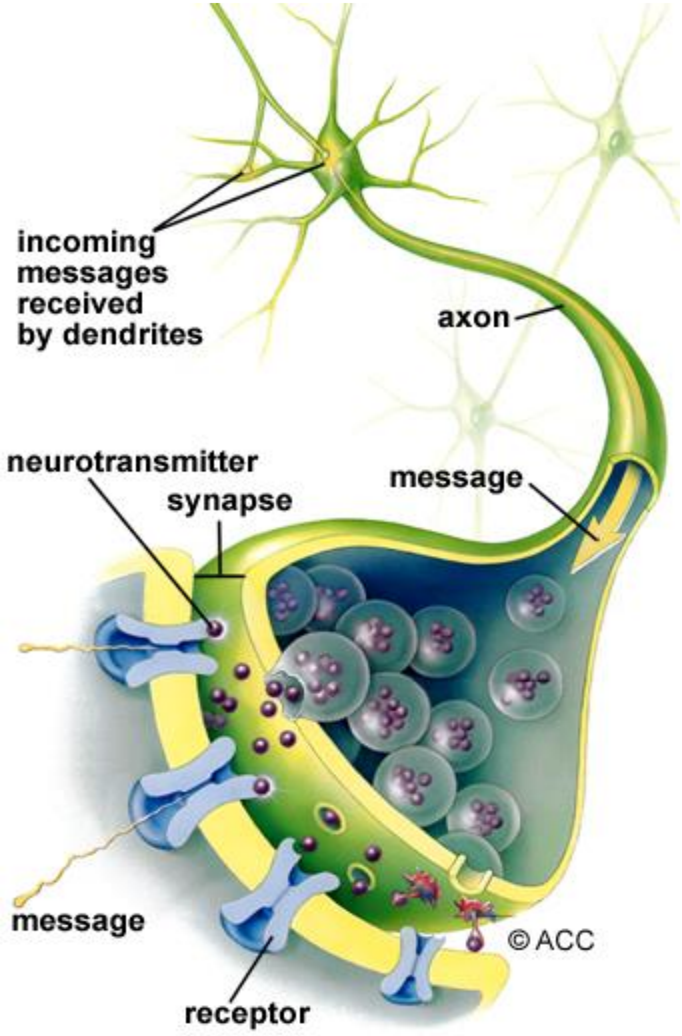
Introduction of Autonomy

- Makes it harder to ensure performance of intended functionality
- Operating conditions harder to quantify
  - Sensor degradation
  - Subsystem malfunctions
  - Operator errors
- Added complexity make interactions harder to constrain

Trust in Automation

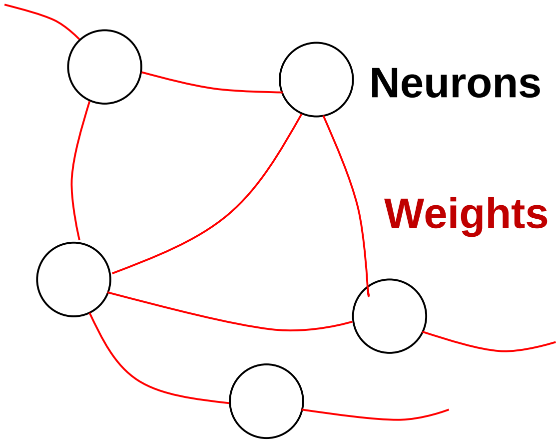
- Current approach to Software:
  - Lots of experience over many years
  - Very conservative design and implementation
  - Established guidelines understood well
  - Prescriptive approach (everyone knows what to do)
  - Verification - Completion criteria understood
- Certification of Autonomy hard
  - Hard to scale up
  - Data in ANNs is unstructured
  - When are we done with testing?

A Neuron and its connections

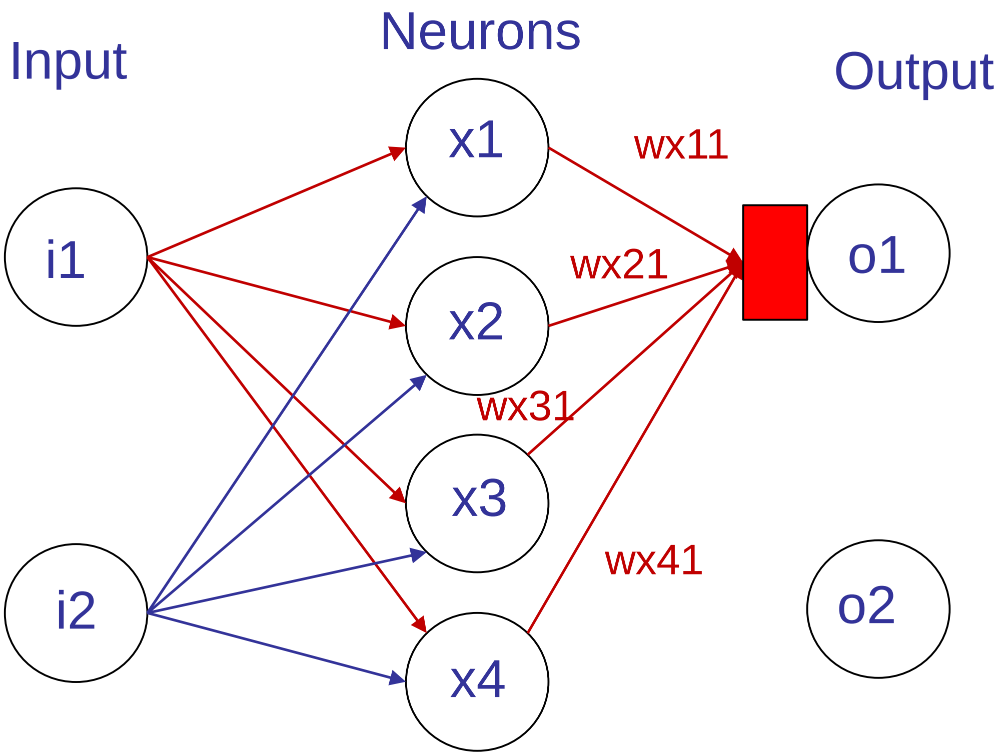


Building Blocks of a Brain

Simplified Representation



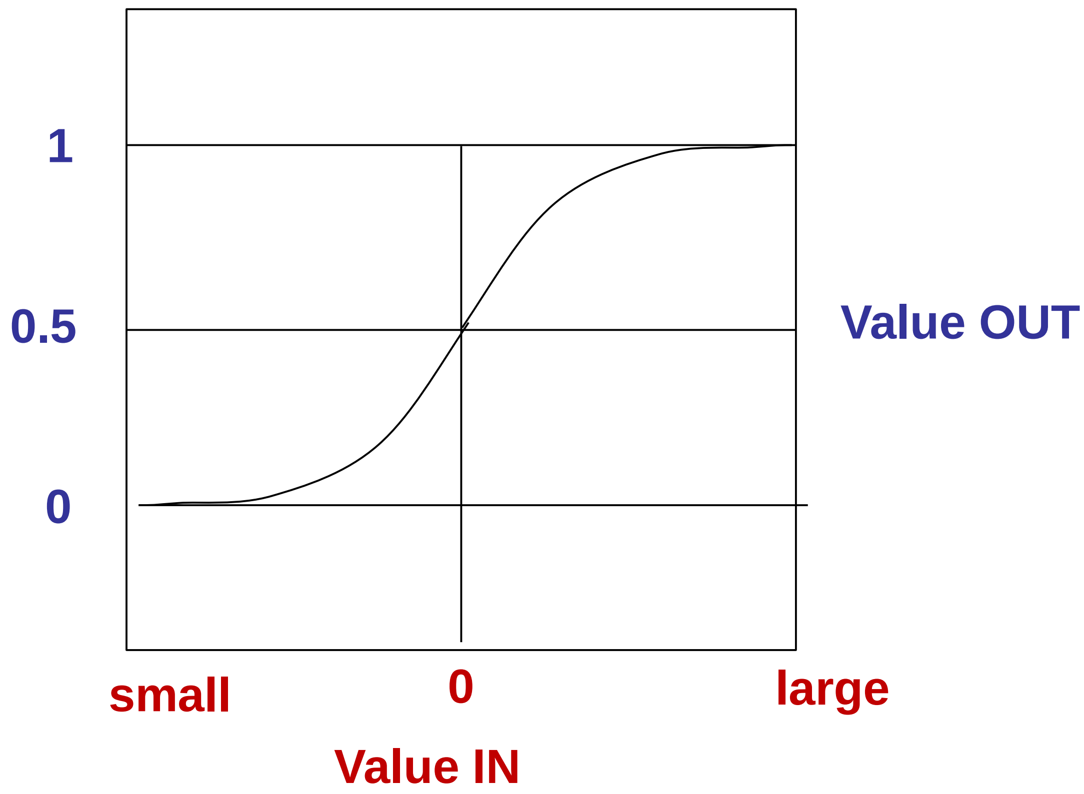
Artificial Neural Network (with Activation)



o1 = Activation (x1*wx11 + x2*wx21 + x3*wx31 + x4*wx41)

Activation Function - example

Sigmoid Function $\frac{1}{1+\ e^{-x}}$



Other functions:

Tanh (x)

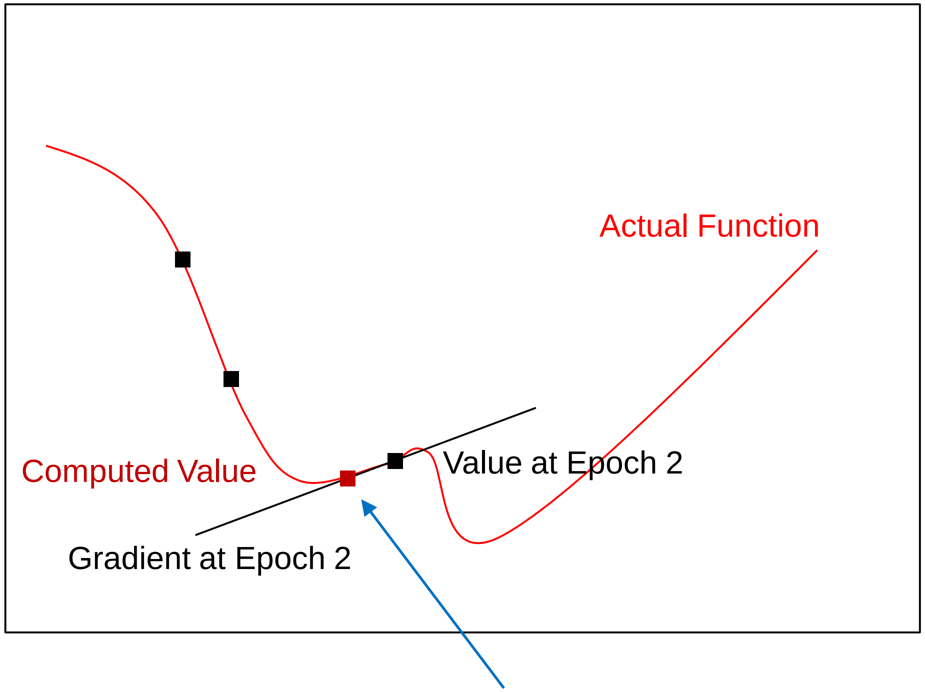
ReLU max(0,x)

And many others…

Reward function using Gradient Descent

wx11 at t=1

wx11 at t=2



Don’t get stuck in the local minima

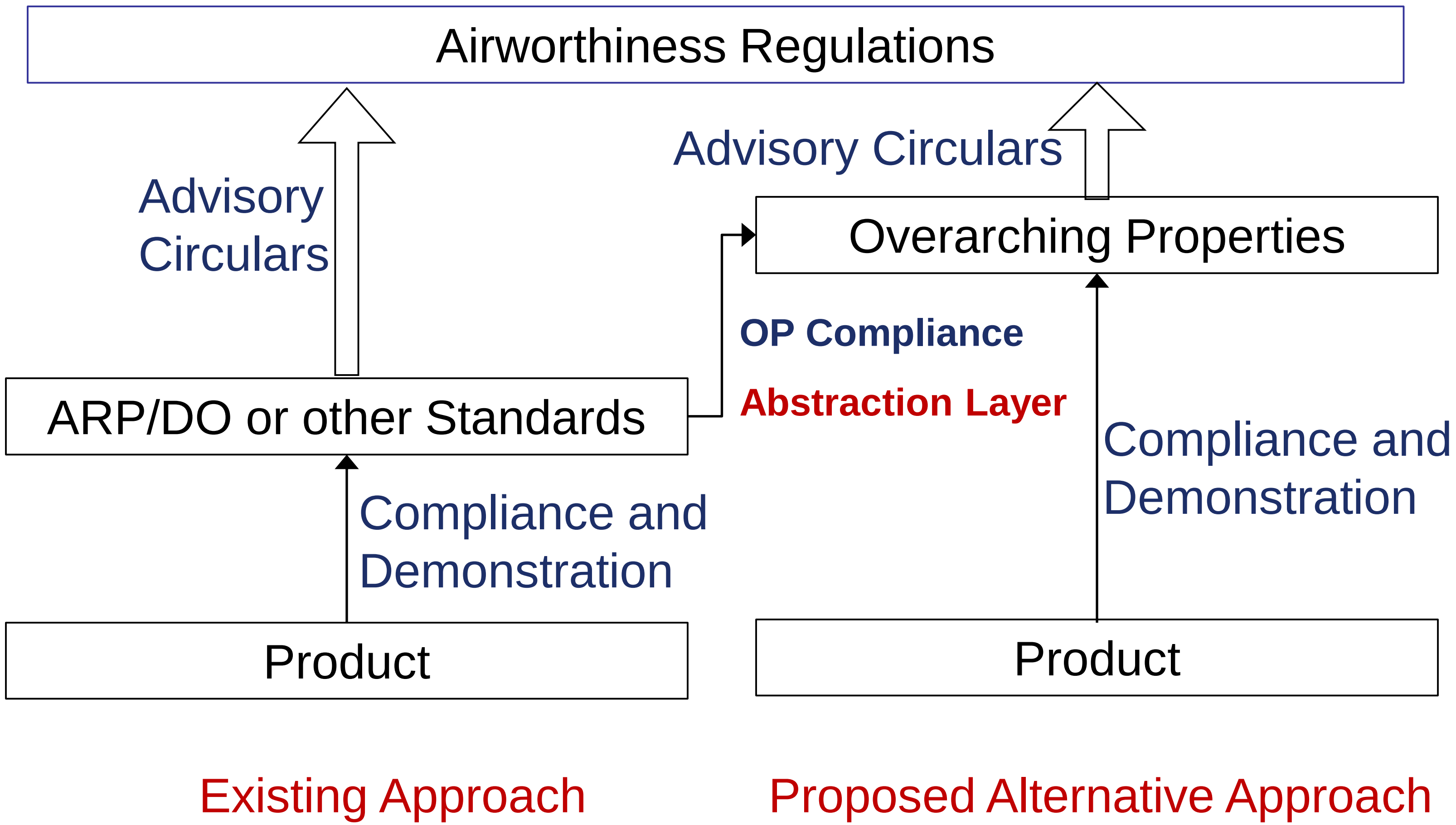
Verification of Artificial Neural Networks

- The algorithms are (typically) straight forward
  - Simple code repeated for all data nodes
  - Code can be verified using customary (DO-178) processes
  - Single set of data vectors could provide coverage over entire code – But!
- The Learned Weights used to perform the Input to Output transformation are hard to verify.
    - No direct correspondence to the expected behavior.
    - Computed by the learning process
- DO-178 does not support verification of an ANN

US Federal Aviation Regulations

- Parts 23 (General Aviation), Part 25 (Transport), Part 27 (Rotorcraft), Part 29 (Transport Category Rotorcraft)…
- “The equipment, systems, and installations must be designed and installed to ensure they perform their intended functions under all foreseeable operating conditions”

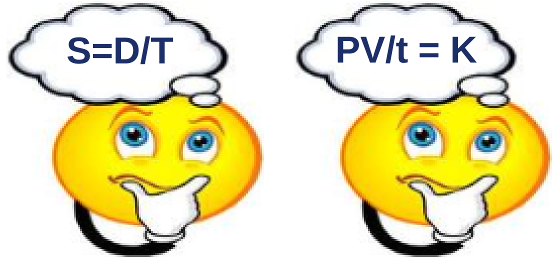
Gaining Approval



Overarching Properties

Stakeholder Needs

- What we think
  we want !
- Intended Behavior,
- Requirements

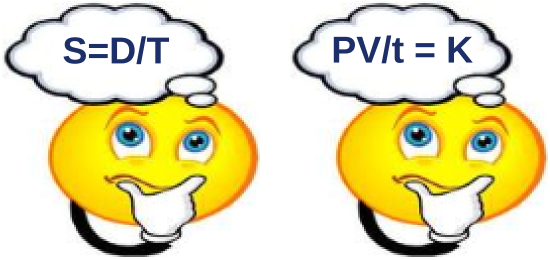


- Customers
- Subject Matter
  Experts
- Users

Overarching Properties

Stakeholder Needs

- What we think we want !
- Intended Behavior,
- Requirements



- Customers
- Subject Matter Experts
- Users

Recorded

Reviewed
Validated

1

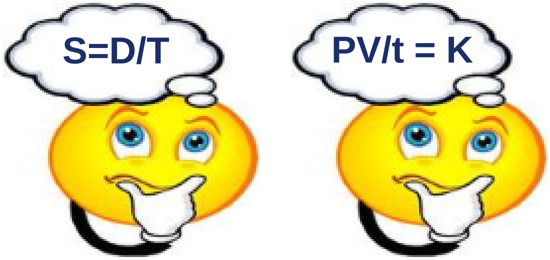
Defined Intended Behavior

Intent Property

Overarching Properties

Stakeholder Needs

- What we think we want !
- Intended Behavior,
- Requirements



- Customers
- Subject Matter Experts
- Users

Recorded

Reviewed
Validated

1 Defined Intended Behavior

Intent Property

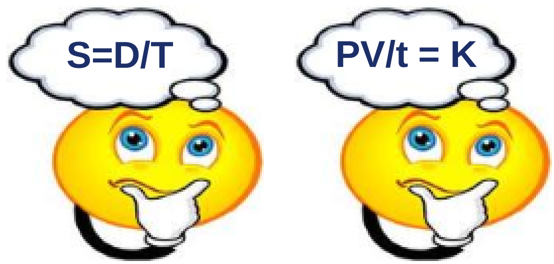
2 Correct Implementation

Correctness Property

Overarching Properties

Stakeholder Needs

- What we think
  we want !
- Intended Behavior,
- Requirements



- Customers
- Subject Matter
  Experts
- Users

Recorded

Reviewed
Validated

1

Defined Intended Behavior

Intent Property

- **No Extraneous Behavior**
- **Or if present, then it does not compromise safety**

3

2

Correct Implementation

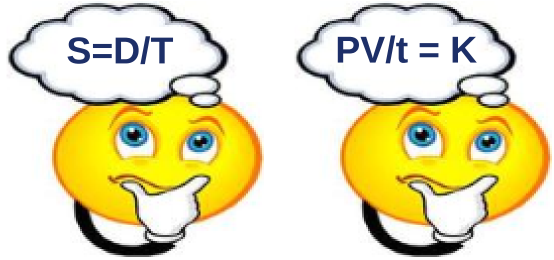
Innocuity Property

Correctness Property

Overarching Properties

Stakeholder Needs

- What we think we want !
- Intended Behavior,
- Requirements



- Customers
- Subject Matter Experts
- Users

Recorded

Reviewed
Validated

1 — Defined Intended Behavior

Intent Property

No Implied Order

- **No Extraneous Behavior**
- **Or if present, then it does not compromise safety**

3 — Innocuity Property

2 — Correct Implementation

Correctness Property

How to show Product “owns” the properties

- Build Assurance Case
  - Communicates a line of reasoning which ties the ownership of the OPs to evidence
  - Should be a structured, compelling argument
- Many notations exist
  - Goal Structuring Notation (GSN)
  - Toulmin
  - Etc.
- Structured Text proposed
  - Can be manipulated by tools
  - Can be translated to graphical forms

Templates and Evidence Schemes

- Developing an approach to produce Assurance Case Templates
- Template Catalog
  - Will help Assurance case adoption
  - Lower cost of certification through reuse

Note!

Assurance Case Templates will help with Understanding the Argument

Verification evidence still required (e.g. Testing)

“OP” Positions are not fixed - yet

- Some
  - Looking to offer more flexibility for applicants
  - Use of Risk based process adjustments
- Other
  - Concerns with applicants having more flexibility:
    - Lack of approval uniformity
    - Hard to educate auditors to reach consistent approval
    - Cannot reach legal approval obligations

Still a work in Progress

Deep Neural Networks

Learning process depends on reward heuristics – (varies with time)

- If learning is continues during operational use, then
  - May not know what to expect
  - Behavior is not uniform
  - Behavior is not under configuration control
  - Cannot show absence of unintended behavior
  - Cannot perform accident investigation
- Learning should be disabled when complete
  - Resource use becomes constant
  - Compute time becomes more predictable (depending on activation trigger optimization)
- Network can be 'tuned' to balance between Resource use, Time and Precision

Bounding Behavior

- Use “Safety Nets” around non-deterministic part of system
- Multiple monitors possible (with voting?)
- Safe Reinforcement learning
  - “Shielding” reward function, teaches only safe actions

Compare Pilot and Artificial Neural Network

- **Training required**
- **Learning through experience is ongoing**
- **Trusted by public**

- **Training required**
- **Learning switched OFF before deployment**
- **Trust not established yet**

If we look inside at the Neurons and connections – we still cannot work out what they are “thinking”

Current Challenge:

how to ensure enough Pilots

how to establish Trust

Proposed uses

- Autonomous – co-pilot
- UAS landing
  - Clear runway
  - Package delivery
- Sense and avoid
- Terrain recognition (follow pipeline)

- Algorithms with discontinuities

Examples of AI/ML in Aviation

- ACAS-Xu - Detect and Avoid System
  - Developed by MIT / Stanford
  - Uses reLUPlex (ANN and Linear Programming)
  - Works well, but not certified (don’t know how)
- Fuel measurement system
  - BF Goodrich
  - Works well, but not certified (don’t know how)

Design Assurance Levels

- Tied to Risk through ARP-4761
  - Catastrophic – Level A
  - Major – Level B
  - Minor – Level C
- No scientific Foundation (best practice approach)
  - How to tie this to AI?
  - It's an economic driving factor –
    - Otherwise just use DAL A.

Research Continues

- ReLUPlex example – Simple activation function, Linear programming constraints (Simplex) ACAS-Xu
- Fuel Measurement example
- For object recognition ANNs may perform better than people – now!
- Automated verification techniques sometimes fail
- Avoiding latent bias (e.g. Wolves and huskies, Stop sign with post-it-note)
- How do we adjust "Leveling"? (DAL A, B, C, D)

Trust in Automation

- Current approach to Software:
  - Lots of experience over many years
  - Very conservative design and implementation
  - Established guidelines understood well
  - Prescriptive approach (everyone knows what to do)
  - Verification - Completion criteria understood
- Makes Certification of Autonomy hard
  - Hard to scale up
  - Data in ANNs is unstructured
  - When are we done with testing?

Research is underway!