CONSIDERATIONS FOR EVOLUTIONARY QUALIFICATION OF SAFETY-CRITICAL SYSTEMS WITH AI-BASED COMPONENTS

Pr. François Terrier, Dr. Huascar Espinoza Ortiz, Dr. Morayo Adedjouma
Département d’Ingénierie des Logiciels et des Systèmes
Software and System Engineering Department
CEA's MAIN MISSIONS

INSTITUTE FOR INTEGRATION OF SYSTEMS & TECHNOLOGIES

AI activities on three axes

AI Tooling and methods

- Safety critical system design
- Expert system platform: Fuzzy, Spatial, Temporal
- Proved embedded constraint solver
- Distributed consensus

AI Deployment technologies

- DNNs configuration & HW mapping
- DNNs accelerator
- Synaptic based
- Three 3D integration

AI applications

- Manufacturing: control by vision
- People, vehicle recognition
- Autonomous vehicle
- Health expert system
- Virtual assistant
- Semantic analysis of multimedia documents

AI Tooling and methods

- Manufacturing default prediction
- Gaz pipe maintenance expert system
- SHM parameter learning

Cybersecurity

- Safety

Industry competitiveness

- 16,000 people
- 600 Industry partners
- 750 Patents/year

Scientific excellence

- Defense and national security
- Energy independance

Industrial technologies

- DRF Fundamental research
- DAM Security-defense
- DEN Civil nuclear energy
- Technological research

Science and Technology

- 4,800 Scientific publications
- 750 Patents/year

Industry competitiveness

- 16,000 people
- 600 Industry partners
- 750 Patents/year

Scientific excellence

- Defense and national security
- Energy independance

Industry technologies

- DRF Fundamental research
- DAM Security-defense
- DEN Civil nuclear energy
- Technological research

Science and Technology

- 4,800 Scientific publications
EXAMPLE OF DEEP LEARNING APPLICATION: REAL TIME VIDEO INTERPRETATION

Vehicle environment interactions understanding

Presented at CES 2018, 2019 on Drive4U stand of Valeo

DeepManta
A many-task deep neural network for visual object recognition

For detection and estimation of orientation and distance

Performances:
+30% / State of the Art
Certification/qualification of safety-critical systems with AI-based components
CURRENT QUALIFICATION PRACTICE FOR SAFETY-CRITICAL SYSTEMS

- Based on (prescriptive) industrial Standards
  - Conformance requirements
  - Prescribe a set of practices
  - Trace the decision, and assessment artefacts

  ⇒ Safety integrity levels

- Assurance effort is proportional to function/system criticality/integrity levels
  - Degree of risk wrt. criticality of failures
  - Process & techniques adapted to each level
  - Highest level applied to most severe failures

  ⇒ Validation costs can add e.g. 30-150% (DO-178)

- Incremental qualification is a core concern
  - Integrated Modular Avionics (IMA) in DO-297
  - Safety Element out of Context (SEooC) in ISO 26262 (Automotive)
  - Generic Safety Case in EN 50129 (Railway)

  ⇒ Modular architectures
...SO, WHY WONDERING ABOUT AI TECHNOLOGY QUALIFICATION?

The technology comes with a short "maturity", with clear weakness on Usage, specification, design, robustness, security...

... DEVELOPMENT PROCESSES ARE (still) NOT UNDER CONTROL!

Formal, traced & rational approach
- Requirements
  - Functions & Sub-functions & Actions
- Each instruction, each value is deduced and justified

An empirical approach
- Informal requirements "by examples"
  - trails and errors
- Poor justification, explanation of the result

AI/ML qualification is still an open issue*
- Unformal requirement with less / no structuration, dynamic evolution of system definition
  - Hard-to-scale-up operating conditions, Verification completion criteria: when are we done with testing?

Breaks all the conformity assessment principles and processes...

*Some standardization bodies are working on AI, e.g. ISO/IEC JTC 1 Standards Committee on Artificial Intelligence (SC 42), EUROCAE WG-114, and the working groups of IEEE SA's AI standards series

Its true for both knowledge based AI and data based AI

With a very pregnant pressure on ML based AI

... to a 3rd AI Winter?
• The frequency of changes is potentially large.
  • AI-based systems are more influenced by obsolescence of data, system's operating environment, sensors…
  ➞ …which leads to need of repetitive/continuous (re-)qualification processes.

• The complexity of the validation process
  • …the costs of revalidation, even for small changes are very high
    e.g. we could need re-training the system
    for slightest modification of a function
    (E.g. deep learning algorithms containing millions of parameters in close interaction)

Re-qualification is easier if the system has been designed with this objective…
Need of Paradigm Change from Different Perspectives:

A- Modularity and Metrics for AI-based System Architectures

B- Evolution: a continuum between Development and Operational Phases
INCREMENTAL CERTIFICATION: BASED ON MODULARITY

- **Principle of Incremental Certification (e.g.: IMA)**
  - Functional component with a well defined perimeter of evolution and consistent with operational specs.
  - Functional Increment (addition or suppression)
  - Enable acquisition of qualification credits at component level

- **Current AI-Based Architectures**
  - Modular architecture with mixed AI/non-AI functions
    - **Pros:** easier to validate, optimize, deploy.
    - **Cons:** error accumulation, calibration is harder
  - End-to-end ANN-based architectures
    - **Pros:** optimal representation wrt to desired task
    - **Cons:** large amount of data, hard to scale, less explainability.
  - Combination Modular/End-to-end
    - Encapsulate driving policy and transfer driving policies from simulation to reality via modularity and abstraction.

---


Driving Policy Transfer via Modularity and Abstraction”, M Muller, A Dosovitsky, et al. 2018.
TRENDS TO HELP SAFETY MANAGEMENT THROUGH MODULARITY

Decomposition of safety-related properties

- Safety can be improved by quantifying component output uncertainties & propagating them forward through the pipeline.
- This improves interpretability by explaining what the different modules observes and why the whole system makes the decisions.

Measure/Prediction of accuracy/uncertainties/probabilities is a key

Contract based approaches

- Demonstrate each function/module meets its safety guarantees under all conditions where the assumptions hold.
- Particularly challenging is finding the boundaries (formal verification can help here!), measuring abnormal situations (including uncertainty) and managing global safety integrity.

Modular approaches and more precise means to specify/measure safe behavior at component level, would help!
Need of Paradigm Change from Different Perspectives:

A- Modularity and metrics for AI-based System Architecture

B- Evolution: a continuum between Development and Operational Phases
Even in the conservative case where we disable learning-based adaptation before deployment we need to:

- Track breaking changes that trigger re-qualification process
- Monitor unacceptable operational conditions (unpredicted)
- Observe inconsistencies between training and operation ML statistical quality.

ML-based systems require pervasive monitoring!
NEED FOR EVOLUTIONARY PROCESSES FOR AI ASSURANCE

Towards an evolutionary AI-oriented lifecycle:

- **Allow each stage** of development and assurance **preserve** the evidence chain in a disciplined way and leaving **auditable records**.
- This needs an explicit specification of the assurance and qualification process, as well as the management of **specific metrics**.
- **Incremental construction**, systematic reviews.

Need to embed **evolution models** in the architecture

- Containing specific principles triggering re-qualification needs.
- Metrics to assess and filter new stimuli and situations (e.g. unpredicted environment conditions).
- Mechanisms to integrate new filtered knowledge so as to grow up system and environment models.

→ We need more integrated development-operation processes and system evolution records to warn qualification stakeholders!
Need of Paradigm Change from Different Perspectives:

A- Modularity and metrics for AI-based System Architecture

B- Evolution: a continuum between Development and Operational Phases

C- Dynamic assurance case
   → metrics definition and process support to build the arguments to justify confidence to the stakeholders (authorities, regulators…) that the system is enough safe, dependable, performant, etc. for the purpose it has been built.
EVOLUTIONARY AI* QUALIFICATION CHALLENGES  
*Focus on ML

And now?
A NEW TOP LEVEL CEA' STRATEGIC PROGRAM for AI TRUST

TOOL & TECHNOLOGY platform for R&D on CERTIFIED AI

Understand

Deploy

Certify

Reinvent learning: Machine discovering


M. E. A. Seddik, M. Tamaazousti and R. Couillet
"Kernel Random Matrices of Large Concentrated Data: The Example of GAN-generated Images." ICASSP’19

Explainable initial model

Why3

Embedded code

Proved constraint solver

Formal language to code solvers.

Formal spec. of safety properties

Proved embedded constraint solver

Formal spec. & verification of DNNs

Ongoing R&D on certification, formal validation, uncertainty measurement, monitoring…

« Happy » AI developers

« Happy » AI users
2. ISO 26262-1:2018, Road vehicles — Functional safety
5. Assuring the Machine Learning Lifecycle: Desiderata, Methods, and Challenges, Rob Ashmore, Radu Calinescu, Colin Paterson