Automated Rapid Certification Of Software (ARCOS)

Ray Richards
Program Manager

Proposers’ Day

14 May 2019

DARPA
<table>
<thead>
<tr>
<th>Time</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:30 AM</td>
<td>Registration</td>
</tr>
<tr>
<td>9:30 AM</td>
<td>Welcoming Remarks, Dr. Ray Richards</td>
</tr>
<tr>
<td>9:35 AM</td>
<td>Security Briefing, DARPA Security</td>
</tr>
<tr>
<td>9:40 AM</td>
<td>ARCOS Program Briefing, Dr. Ray Richards</td>
</tr>
<tr>
<td>10:00 AM</td>
<td>Software Certification, Dr. John Seel</td>
</tr>
<tr>
<td>10:20 AM</td>
<td>Seedling Results, Dr. Gustavo Quiros Araya, Siemens Corporation</td>
</tr>
<tr>
<td>10:40 AM</td>
<td>Break and Submit Questions</td>
</tr>
<tr>
<td>11:00 AM</td>
<td>Contracts Management Office Briefing, Mr. Mark Jones, DARPA CMO</td>
</tr>
<tr>
<td>11:30 AM</td>
<td>Teaming Intro Briefs - 2 min each** (Part 1)</td>
</tr>
<tr>
<td>12:00 PM</td>
<td>Lunch Break (on your own)</td>
</tr>
<tr>
<td>1:00 PM</td>
<td>Question and Answer Session</td>
</tr>
<tr>
<td>2:00 PM</td>
<td>Teaming Intro Briefs - 2 min each** (Part 2)</td>
</tr>
<tr>
<td>2:30 PM</td>
<td>Networking Session - Conference room will remain available</td>
</tr>
<tr>
<td>3:30 PM</td>
<td>Meeting Adjourns</td>
</tr>
</tbody>
</table>
Automate the evaluation of software assurance evidence to enable certifiers to rapidly determine that the risk of software deployment is acceptable.
Current software certification practices are unsustainable

- Our ability to produce software has outpaced our ability to certify software
- The cost and time needed to certify software is an impediment to fielding new capability
- Software certification cost is a major consideration in managing a systems lifecycle.
- Certification example: Multi-level secure operating system certified for F22 and F35 in accordance to the Common Criteria
  - Previously certified in safety critical avionics
  - Evaluation underway by April 2006
  - Certificate awarded January 2011

DSB Report on Design and Acquisition of Software for Defense Systems, February 2018

Software in military systems continues rapid growth.

Thousands of Lines of Code

- F-16A Block 1 (1974)
- F-16D Block 60 (1984)
- P-3C (1996)
- F-22 Raptor (1997)
- F-35 Lightning II (2006)
- F-35 Lightning II Block 2B
- F-35 Lightning II Block 3F est.

National Information Assurance Partnership

Common Criteria Certificate

Certificate awarded January 2011
Certifications today: Up to 5 additional years post-development to obtain ATO

- REQUIREMENTS
- Design, Build, and Document
- Design Documentation
- Test and Generate Assurance Evidence
- Evaluate Assurance Evidence
- ATO

Evidence:
- Test
- Analysis
- Simulation/Emulation
- Software Quality Assurance (SQA)

Human-driven activities produce implementation and assurance evidence in human readable documents.

Assurance evidence captured in human readable documents.

Manual activity evaluates whether or not criteria are met.

ATO: Authority to Operate
Certifications today: Up to 5 additional years post-development to obtain ATO

Certifications with ARCOS: ATO concurrent with development

Co-generation of assurance evidence with documentation and implementation artifacts

Implementation and assurance evidence maintained in binary formats

Automated generation and validation of assurance case arguments

Distribution A. Approved for public release: distribution unlimited.
ARCOS Program Structure
TA1: Generation of high quality assurance evidence to support curation and facilitate TA3 analytics

TA2: Curate a large disparate body of evidence to support TA3 analytics

TA3: Generate assurance arguments and use data analytics to back the arguments with evidence

TA4: Provide a series of challenges culminating in the processing of evidence and generation of assurance for a realistic system

ARCOS Technical Approach and Technical Areas

**TA1: Evidence Generation**
- **New Software**
- **Legacy Software**

**Evidence Evaluation**
- **TA2: Evidence Curation**
- **TA3: Assurance Generation**

**TA4: Quantitative Assessment**
- **Human**
- **Automated**

Requirements flow through the process of evidence generation, curation, and evaluation, culminating in assurance arguments and quantitative assessment.
**Goal:** Create technologies to construct assurance cases that drive the design and implementation of software

**Challenges:**
- Develop mechanisms to construct assurance cases, connecting them to software development tools
- Automate the generation of evidence to enable construction of assurance cases
- Capture component interface specifications to support composition

**Possible Approaches:**
- Invent assurance case languages and tools
- Automate evidence generation through software analysis and test case generation and use
- Extend architectural languages and tools
Goal: Create technologies to produce assurance evidence to support certification of existing software

Challenges:

- Generate strong evidence that is focused on certification criteria and assurance cases
- Automate the generation of evidence through test, emulated execution, and analysis

Possible Approaches:

- Extend NLP approaches for the extraction of information
- Invent techniques for the discovery of high-level structures in legacy software
- Create mechanisms to derive specifications for the software using machine learning algorithms
- Automate software analysis, test generation and targeted fuzzing to augment evidence
**TA 2: Evidence Curation**

**Goal:** Provide machine readable, common representation of assurance evidence with traceable provenance

- Develop a representation of evidence from disparate domains that is amenable to analysis
- Developing mapping of evidence to system structure
- Attest to provenance of evidence (chain of custody)

**Possible Approaches:**
- Create mechanism to make inference from function to structure
- Scale-up the graph-of-graphs representation explored in seedling effort
- Develop provenance trees to attest to integrity of artifacts
TA 3: Assurance Generation

**Goal:** Create technology to automate generation of assurance cases from curated evidence

**Challenges:**
- Construct sound assurance arguments
- Substantiate arguments with strongest available evidence
- Evaluate assurance cases to determine soundness and confidence score

**Possible Approaches:**
- Scale up automated theorem prover technology to construct design-based assurance case arguments
- Develop trained classifiers identify evidence that best fits arguments
- Extend formal methods to demonstrate soundness and correctness of validation algorithms
**TA 4: Quantitative Assessment**

**Goals:** Provide a testbed and a series of challenges culminating in demonstrations of processing evidence and generation of assurance for a realistic system. Perform periodic technology assessments.

**Challenges:**

- Perform quantitative assessment of ARCOS technologies in coordination with Government evaluators
- Provide a progression of increasingly challenging software systems and data sets that are accessible to research performers
- Demonstrate ARCOS technologies to a military relevant software system in the final phase
- Develop a conservative and sound approach to sample data for evaluation
• Provide evaluations for quantitative assessments
  • Evaluate generated evidence and assurance cases
  • Identify and document shortcomings
• End of each phase report on applicability to government C&A practices
• During each phase
  • Provide feedback on the strengths and weakness of performers
  • Compare proposed approaches with the current state of practice
• Team members will include certification experts from across the DoD and Federal Government
## ARCOS Goals and Metrics

<table>
<thead>
<tr>
<th></th>
<th>Phase 1 (18 Months)</th>
<th>Phase 2 (18 months)</th>
<th>Phase 3 (12 Months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale</td>
<td>One module from Phase 3 system</td>
<td>Set of modules from Phase 3 system</td>
<td>Realistic System</td>
</tr>
<tr>
<td>Evidence Domains</td>
<td>Test, Simulation, and Emulation</td>
<td>Test, Simulation, Emulation, Analysis, and SQA</td>
<td>Test, Simulation, Emulation, Analysis, and SQA</td>
</tr>
<tr>
<td>Assurance Generation</td>
<td>&lt;4 weeks</td>
<td>&lt;2 weeks</td>
<td>&lt;1 week</td>
</tr>
<tr>
<td>TA 1 Legacy Software</td>
<td>3 of 6 evaluators agree with generated evidence</td>
<td>4 of 6 evaluators agree with generated evidence</td>
<td>5 of 6 evaluators agree with generated evidence</td>
</tr>
<tr>
<td>TA 1 New Software</td>
<td>3 of 6 evaluators agree with generated evidence</td>
<td>4 of 6 evaluators agree with generated evidence</td>
<td>5 of 6 evaluators agree with generated evidence</td>
</tr>
<tr>
<td>TA 2</td>
<td>100,000+ nodes curated</td>
<td>1,000,000+ nodes curated</td>
<td>10,000,000+ nodes curated</td>
</tr>
<tr>
<td>TA 3</td>
<td>3 of 6 evaluators agree with confidence scores</td>
<td>4 of 6 evaluators agree with confidence scores</td>
<td>5 of 6 evaluators agree with confidence scores</td>
</tr>
<tr>
<td>TA 4</td>
<td>Performs quantitative assessments with Government Evaluation Team</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# Schedule

<table>
<thead>
<tr>
<th>ARCOS</th>
<th>Phase 1: Initial Capability</th>
<th>Phase 2: Complete Capability</th>
<th>Phase 3: Scale to Realistic System</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>18 Months</td>
<td>18 Months</td>
<td>12 Months</td>
</tr>
</tbody>
</table>

## Technical Areas

<table>
<thead>
<tr>
<th>TA 1: Evidence Generation</th>
<th>Component Evidence Generation</th>
<th>Composed Evidence Generation</th>
<th>Scale Evidence Generation</th>
</tr>
</thead>
<tbody>
<tr>
<td>TA 2: Evidence Curation</td>
<td>Curate Test, Simulation, and Emulation Evidence</td>
<td>Curate All Evidence</td>
<td>Scale Evidence Curation</td>
</tr>
<tr>
<td>TA 3: Assurance Generation</td>
<td>Generate Component Assurance</td>
<td>Generated Composed Assurance</td>
<td>Scale Assurance Generation</td>
</tr>
<tr>
<td>TA 4: Quantitative Assessment</td>
<td>Component</td>
<td>Components and Composition</td>
<td>System</td>
</tr>
</tbody>
</table>

Program Meetings
- TA 1 Assessments
- TA 2 Assessments
- TA 2-3 Assessments

Distribution A. Approved for public release: distribution unlimited.
Awards

- One TA per proposal
- TAs 1 & 3
  - Multiple awards anticipated for each TA
  - Responding to a portion of the TA is acceptable
    - TA 1
      - Legacy Software
      - New Software
    - TA 3
      - Assurance Generation
      - Assurance Validation
  - Addressing the entire TA is encouraged
- TA 2 & 4
  - Single award anticipated for each TA
  - TA 4 performer can not have a conflict of interest with other performers
- Collaborative proposals are encouraged
Important Dates

- Proposers’ Day: May 14, 2019
- Abstract Due Date: May 24, 2019, 12:00 noon (EDT)
  - Feedback will be provided as quickly as possible
- Proposals Due / BAA Closing Date: July 9, 2019, 12:00 noon (EDT)