Using evidence-based arguments to support dependability assurance – experience and challenges
(invited presentation)

Janusz Górski
Departament of Software Engineering, Gdańsk University of Technology
jango@pg.edu.pl

Workshop on Challenges and New Approaches for Dependable and Cyber-Physical Systems Engineering (DeCPS)
Warszawa, 14th June 2019, Poland
Janusz Górski

► Professor at Gdańsk University of Technology (GUT)
► Leader of Information Assurance Group (IAG) – a research group at Department of Software Engineering, Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology (http://iag.pg.gda.pl/)
  • Focusing on trust and risk management of computerized systems and services
  • Trust-IT methodology and the NOR-STA tool supporting application of evidence-based arguments to analyse and demonstrate assurance and compliance
► Proxy of ARGEVIDE sp. z o.o. – a spin-off of GUT
► Present involvement
  • EWICS Security (European Workshop on Industrial Computer Systems) MED&SEC
  • ISA99 Committee (International Society of Automation)
  • ICCS/ERNCIP (IACS components Cybersecurity Certification Schemes)
  • IoTSec (Internet of Things Security)
IAG R&D related to evidence-based arguments and their applications

► Research
  • 1993–1995 Project SHIP (Safety of Hazardous Industrial Processes), European Program ENVIRONMENT
  • 2001 – 2003 Project DRIVE (Drugs in Virtual Enterprise), 5th EU FR
  • 2004 – 2007 Project PIPS (Personalised Information Platform for health and life Services), EU 6th FR
  • 2006 – 2008 Project ANGEL (Advanced Networked Gateway to Enhance quality of Life), EU 6th FR
  • 2009 Project ERM (Selected Problems in Environmental Risk Management and Emerging Threats), Polish-Norwegian Research Fund

► Industrial trials
  • 2010-2014 – Project NOR-STA, European Regional Development Fund

► Commercialization
  • 2014- now, ARGEVIDE sp. z o.o. – a spin-off of GUT (www.argevide.com)
    ▶ Customers in Oil&Gas, Medical, Maritime, Railways, Automotive and Aviation sectors

Acknowledgement (significant contributors to this R&D)
Dr Łukasz Cyra, Jakub Czyżnikiewicz (programmer),
Dr Aleksander Jarzębowicz, Dr Jakub Miler,
Dr Andrzej Wardziński, Michał Witkowicz (programmer)
Contents

• What are the evidence based arguments and what are they for?
• Selected challenges
  – Argument representation
  – Support for communication and co-operation
  – Argument assessment
  – Scalability and change management
  – Integration
  – Argument structuring and reuse
  – Composability
  – ’Living’ arguments
• Experiences from developing and deploying a tool for supporting evidence based arguments
Evidence-based argument
What is it?
Evidence-based arguments

- **Argument** is an attempt to persuade someone of something, by giving reasons and/or evidence for accepting a particular conclusion.

- This ‘something’ can be:
  - assurance of some important property (safety, security, privacy, reliability, …)
  - conformance with a stated set of criteria (standard, norm, directive, recommendation and so on)
  - ...

- **EXAMPLE ARGUMENT**

  Tests confirm that this software module satisfies its requirements because test results are positive and the test coverage is sufficient.
Evidence-based arguments

- **Evidence** in its broadest sense *includes everything that is used to determine or demonstrate the truth of an assertion.*
  - Evidence can be used in arguments – to demonstrate the truth of the premises

EXAMPLE ASSERTION: It is raining outside

EVIDENCE:

Assumption:
Evidence is delivered in electronic documents of any form: text, graphics, image, video, audio etc.
  - *.txt, *.doc *.xls *.jpg *.mp3, *.pdf, *.mp4, ...
Tests confirm that this software module satisfies its requirements because test results are positive and the test coverage is sufficient.

**Claim:** Module meets requirements

**Strategy of argumentation:** Argumentation by referring to test results and test coverage

**Rationale:** Experience shows that positive results of tests of adequate coverage reliably demonstrate fulfillment of the requirements

**Evidence:**
Demonstrates the fact about test results and test coverage

**Fact:**
- Adequate coverage and positive test results
A case study: Evidence based argument about quality of a software module

Claim: Module meets requirements

Strategy of argumentation and its rationale

Fact: Adequate coverage and positive test results

Assumption: Adequate configuration control in place

Claim: Adequate competencies of testers

Evidence
Evidence-based arguments
What are they for?
Argument and trust

Convincing arguments can be used to build trust

➤ because they demonstrate trustworthiness

Such arguments we call *Trust Cases*

Example:

A convincing (supported by evidence) argument that a service is secure increases trust in the service

*Evidence:*

*protective measures used,*
*certification procedures passed,*
*penetration tests results,*
*operating data,*
*development practices used …*
Different types of trust cases

Assurance Cases
safety, security, privacy, dependability, reliability ...

Conformance Cases
standards, norms, directives, regulations ...

Metaphysical Cases
e.g arguing the existence of Santa Claus

and others...
Trust-IT and NOR-STA
Trust cases

Evidence based arguments
Deployment (SaaS)

Application specific packages

Generic Argument Management Services

NOR-STA Core

API
(Selected) challenges and experiences
Represenatation
Trust Case Language (TCL)
and the underpinning argument model
Claim

Argumentation strategy

Assumption

Fact

Reference

Information

Evidence

Premises

Inference

Rationale

Conclusion
A case study: Evidence based argument about quality of a software module

Tests confirm that this software module satisfies its requirements because test results are positive and the test coverage is sufficient.

With the assumption that this module was not changed during testing, the tests performed by competent testers confirm that the module satisfies its requirements because test results are positive and the test coverage is sufficient.
Communication and co-operation
Communication and co-operation

- Argument sharing
- Multiple viewpoints (managers, suppliers, certifying/qualifying institutions, argumentation developers, evidence suppliers, external world, ...)
  - Different roles (developer, assessor, viewer, administrator, ...)
  - Access control
  - Different views at the argument
- Support for decision making
  - Argument assessment
- Support for consensus building
- Support for disputes
Argument assessment
Assessing the ‘compelling power’ of argument
Tests confirm that this software module satisfies its requirements because tests results are positive and test coverage is sufficient

Logic doubt: Do successful tests of right coverage really determine the success of testing?

Epistemic doubt: Do we really have positive test results and the right coverage?
The assessment process

Assess conclusions

Assess inferences (local)
Assess facts (local)

For argumentation scheme
For a concrete argument

Example argument:
- Claim 1
  - Argumentation Strategy 1
    - Rationale 1
    - Fact 1
    - Fact 2
  - Fact 3
  - Claim 3
  - Claim 4
- Claim 2
  - Argumentation Strategy 2
    - Rationale 2
    - Fact 3
    - Claim 5
  - Fact 4
  - Claim 6
    - Argumentation Strategy 3
      - Rationale 3
      - Fact 5
- Fact 7
Can we automatically aggregate the local assessments (of inferences and facts) into the assessment of the whole argument?
Assessment of an argument
(based on Dempster-Shafer believe functions)

Assessment of evidence

- Fact: ‘test results are positive’
  
  Test report for this module demonstrating that the test results are positive
  Test report for different module
  Test report for this module demonstrating that the tests failed

- Assessment
  
<table>
<thead>
<tr>
<th>Acceptance</th>
<th>Uncertainty</th>
<th>Rejection</th>
</tr>
</thead>
</table>

Assessment of an argument
(based on Dempster-Shafer believe functions)

Assessment of evidence

– Fact: ‘test results are positive’
  Test report for this module demonstrating that the test results are positive
  Test report for different module
  Test report for this module demonstrating that the tests failed

– Assessment

<table>
<thead>
<tr>
<th>Acceptance</th>
<th>Uncertainty</th>
<th>Rejection</th>
</tr>
</thead>
</table>

Assessment of inference

– ‘if we have positive test results and adequate tests coverage, then the module meets its requirements’
  How reliable is such reasoning?

– Assessment

<table>
<thead>
<tr>
<th>Acceptance</th>
<th>Uncertainty</th>
<th>Rejection</th>
</tr>
</thead>
</table>
Linguistic values make the scale more human friendly:

**Decision**: rejectable, opposable, tolerable, acceptable

**Confidence**: sure, very high, high, low, very low, uncertain
Communicating the assessment results

- Evidence based argument about quality of a software module
  - Module meets requirements
  - Argumentation by referring to test results, test coverage and testers' competencies
  - Experience shows that positive results of tests of adequate coverage if performed by competent testers, reliably demonstrate fulfillment of the requirements
  - Tests results are positive
    - Report from testing
  - Test coverage is satisfactory
    - Test plan
    - Analysis of the test plan from the test coverage viewpoint
  - Adequate configuration control in place
  - Adequate competencies of testers

- Evidence based argument about quality of a software module
  - Module meets requirements
  - Argumentation by referring to test results, test coverage and testers' competencies
  - Experience shows that positive results of tests of adequate coverage if performed by competent testers, reliably demonstrate fulfillment of the requirements
  - Tests results are positive
    - Report from testing
  - Test coverage is satisfactory
    - Test plan
    - Analysis of the test plan from the test coverage viewpoint
  - Adequate configuration control in place
  - Adequate competencies of testers
Presently 9 different methods of argument assessment are implemented:

- Dempster-Shafer
- ISO 33000 (SPICE, Automotive SPICE, ...)
- Rating scale (numerical)
- Three-level assessment
- and others...
Scalability and change management
Operating large arguments

• Large arguments are difficult to handle and to understand
  – What does it mean ‘large’?
    • Experience with arguments up to 8000 nodes

• Graphical representation inadequate
  – Adding/modifying a node can change the graph in two dimensions
  – Adding more explanatory text expands a node and gives a false feeling of growing importance of the node
Operating large arguments

- Large arguments are difficult to handle and to understand. What does 'large' mean?

- Experience with arguments up to 8000 nodes:
  - Graphical representation inadequate:
    - Adding/modifying a node can change the graph in two dimensions.
    - Adding more explanatory text expands a node and gives a false feeling of growing importance.
Managing massive evidence

- Integrating any electronic document as an evidence container: text, graphics, video stream, audio, ...
- Providing for referencing any place the document is stored (web pages, ftp, svn, ...
- Referencing selected fragments of bigger documents (pages, chapters, sections, ...
- Providing for user selected repositories

Open PCA Pump Assurance Case

An argument that Kansas State University's Open PCA Pump design is both acceptably safe and effective

Subject of Assurance Case: PCA Pump

Requirements: Draft 0.11

Background Information

'Major' Level of Concern

External Infusion Pumps are FDA Class II Devices

Claim 0: PCA pump is effective in its medical function and is acceptably safe

Strategy 0: Argue for safety and effectiveness separately, but coordinated

Rationale 0: No medical device can be completely safe; its benefit must justify its risk

Claim 1: PCA pump is effective

Strategy 1: PCA pump performs intended function which has been clinically verified

Rationale 1: PCA pump must perform intended function; that function must be medically necessary

Claim 1.1: PCA pump performs intended function

Divide into individual behaviors, that they are performed correctly, and their composition has inter

4 PCA Pump Function

4.1 Basal Flow Rate

The PCA pump infuses at prescribed basal, bolus, or KVO rates.

4.2 Patient-Requested Bolus

Upon patient's press of the PCA pump's patient-button, a prescribed bolus volume to be infused, $V_{Bolus}$, is delivered in the pump is delivered to the patient. $F_{Bolus}$ is the flow rate of the pump, $F_{Bolus}$, in addition to the prescribed basal flow rate, $F_{Bolus}$, but no more than the maximum flow rate for the pump, $F_{max}$.
Managing multiple arguments and multiple users

- User accounts: roles and permissions

- Containers for arguments: folders and projects
  - Project contains a single argument
  - It is useful to group them together in folders to facilitate access and to enforce common policies
    - Patterns libraries
    - Assurance cases of subsystems
    - Arguments dedicated to a specific user
    - ...

• Manager
• Auditor
• Editor
• Viewer
• Developer
Managing multiple arguments and multiple users

- **User accounts**: roles and permissions
- **Containers for arguments**: folders and projects – a project contains a single argument – it is useful to group them together in folders to facilitate access and to enforce common policies
- **Patterns libraries**
- **Assurance cases of subsystems**
- **Arguments dedicated to a specific user**

[Image of a user interface showing the organization of folders and projects]
Change control

- Evolution of argument, assessments and evidence
- Baselines
- Rollback
- Accountability of changes
Reporting

- Customizable Excel reports
  - assessment history

- Customizable XML/HTML reports
  - XLS scripts to process XML data
  - assessment history

- Dedicated reports
  - Project metrics
  - Project change list
  - NOR-STA users' activities (for administrators)

- GSN diagrams generated for argument sections
Integration
Integration

- Evidence
- External systems
- SACM (Structured Assurance Case Matamodel)
Integrating arguments with evidence

- **Direct links** to evidence (resources on the internet)

- **Internal** NOR-STA repository

- **External repositories** (HTTP Basic Authentication, for example SVN, GIT)

- External repository with **webservice** interface for listing documents (used for integration with Siemens Teamcenter)
Integration with other systems

NOR-STA API (webservices)
• JSON REST webservices cover full NOR-STA functionality

Single Sign On (SSO)
• Active Directory Federation Services ADFS (Oauth 2.0)
• Azure B2C

XML export/import
• TCL format

SACM 2.0 compliance
• NOR-STA use TCL (Trust Case Language) notation which complies to SACM 2.0 (published March 2018)
Argument structuring and reuse
Argument structuring and reuse

– Links in the argument structure
  • DAG instead of tree
– Patterns and pattern libraries
– Templates
  • Deriving structure from standards
  • Following changes in standards
– Deriving argumentation structure from models
  • Architectural models
  • Risk analysis reports
Argument templates and changes propagation
Composability
Composability – modularisation of arguments
• Interfaces between the components and the embedding argument
• Changing context can invalidate evidence and argumentation strategies
Simple solution: syntactic match *Required interface, Provided interface*

Difficult problem: semantic contracts with change control
’Living’ arguments
Living argument

S&P&D Objectives

Logic decomposition into more specific objectives

Analytical and measurement layer – collecting evidence that demonstrates objectives

Smart grid
Living argument

S&P&D

Objectives

Logic decomposition into more specific objectives

Evidence
Evidence
Evidence
Evidence
Evidence

Analytical and measurement layer – collecting evidence that demonstrates objectives

Smart grid
Conclusion

- Argument is a focal point situated between different stakeholders and addressing their important concerns
  - Argument model and its representation – a crucial decision
  - SaaS model of deployment
- Argument is un ‘umbrella’ under which we can integrate the results of a wide range of more focused analytical methods and techniques
- Conformance arguments have a potential to support emerging certification frameworks
  - Cybersecurity of components (and systems)
  - Qualification of medical devices
- Discovering new application domains
  - ‘Customer driven’ development
- For materialisation of the vision of ’living’ arguments more automation is needed
  - automatic determination of an argumentation structure
  - automatic evidence collection and assessment
  - strong context awareness
- SLA, Data Security and Privacy Protection – of high and growing relevance
Thank you for your attention